9-2009

Estimation of Bidder Valuations in an FCC Spectrum Auction

Jungwon YEO
Singapore Management University, jwyeo@smu.edu.sg

Estimation of Bidder Valuations in an FCC Spectrum Auction

Jungwon Yeo *

University of Minnesota

September 1, 2009

Abstract

The Federal Communications Commission (FCC) uses auctions to allocate radio spectrum frequencies to wireless service providers. The innovation of the auction design is that it offers many heterogeneous licenses simultaneously in one ascending auction. This paper develops an empirical model and procedure to estimate bidder valuations. Given that the complex nature of the auction does not admit formal modeling in a general setting, I do not explore a particular model of equilibrium bidding. Instead, I propose two revealed preference inequalities which should hold in any reasonable model of these auctions. The first inequality requires that a bidder never bids on a license at a bidding price above the expected marginal revenue of the license. The second inequality is that if a bidder bids on license A, but not on license B, the expected marginal surplus from winning license A at the bidding price is greater than that of license B. I employ an estimation strategy that generates a map from the observed bidding behavior to a set of distributions of bidder valuations consistent with these behavioral assumptions. A part of the strategy uses an estimator developed by Pakes, Porter, Ho and Ishii (2006). An advantage of this estimator is that it can accommodate a flexible specification of unobserved heterogeneity. It allows bidder-license specific values to capture private information. I apply the empirical model to an auction held in 2006. Using the estimated distribution of bidder valuations, I estimate bidder markups in order to gauge the level of competition in this auction. The estimated bidder markups are large: the median for local bidders such as rural telephone companies is 26%, whereas it is 31% for global bidders such as nationwide carriers. This suggests that there were large distortionary effects of informational rents in the auction.

*I am grateful to my advisor Patrick Bajari for his continuous support and guidance. I also wish to thank Amil Petrin, Kyoo Il Kim, Minjung Park, Tom Holmes and Bob Town, as well as participants of the Applied Micro seminar at the University of Minnesota for their helpful comments. Of course, all remaining errors are my own. Email: yeox0010@umn.edu

1 Introduction

The Federal Communications Commission (FCC) uses auctions to allocate radio spectrum frequencies to wireless service providers. There is no standard method to estimate demand for spectrum licenses because there is no well-accepted model of bidding that captures the complexity of the auction. This paper develops an empirical model and a procedure to estimate bidder valuations and then applies them to a recent auction, labeled Advanced Wireless Service (AWS)-1.

FCC spectrum auctions have attracted much attention from both policy makers and economists for the following reasons. First, spectrum auctions have become an important source of federal government revenue through the more than 60 auctions the FCC has conducted since 1994. For example, two auctions, held in 2006 and 2008, generated a combined $33 billion in revenue for the US Treasury. Second, efficient allocation of spectrum licenses is important because it can promote competition in the highly concentrated wireless telecommunication industry.[1]

The FCC sells many licenses simultaneously in a single ascending auction. A licensee can serve a geographically distinct area using a particular band of spectrum frequencies. An auction proceeds in rounds. During a round, a bidder may submit bids on as many licenses as it wishes. An auction continues until none of the bidders place a new bid on any of the licenses. At that point, the highest bidder for an individual license becomes the winner of the license. This unique auction mechanism, namely the simultaneous ascending (SA) auction, is one of the most frequently cited examples where economic theories are applied to market design.[2]

The simple allocation rule stands at the center of the complex nature of the auction game because it implies that licenses are priced individually although there may be "complementarities" among licenses. In other words, the value of a bundle of licenses may exceed the sum of the values of the individual licenses in the bundle. In this environment, a bidder's willingness-to-pay for a license will depend on the other licenses it will win. This gives rise to a combinatorial property on the bidder's strategy: the bidder needs to consider all possible combinations of licenses it may win when making the decision to bid for a particular license. As a result, economic theory has not been able to characterize an equilibrium in a general setting, and hence provides little guidance on how to interpret bid data generated in an actual auction.

[1] According to the US Wireless Communication Association, the wireless telecommunication industry generated $118 billion in revenue and contributed $92 billion to US GDP in 2004. According to the same source, in 2004, two firms account for 53% of all industry revenue and four firms account for 90%.

[2] Many leading auction theorists were involved in the design process. According to the National Science Foundation, Paul Milgrom, Robert Wilson and Preston McAfee were the main academic contributors to the original FCC spectrum auction design. See Roth (2002), Ausubel and Milgrom (2001), and Bykowsky, Cull and Ledyard (2000) for discussion on the design.

This also makes estimating the distribution of private information on the values of licenses extremely challenging. In the presence of complementarities, the dimension of private information equals the number of possible combinations, $2^N - 1$, where N denotes the number of licenses for sale. Given that many licenses are offered for sale in a single auction, this number can be very large.[3] There is neither an auction theory nor an econometric technique that enables a researcher to explore this path.

As a response to these challenges, I set a moderate goal of estimating only the set containing the true distribution of bidder valuations, while allowing private information to be specific to each bidder and license pair. This goal is achieved through behavioral assumptions and an estimation procedure that generates a mapping from bid data to a set of parameters consistent with the assumptions. In the application to the AWS-1 auction, I test the presence of complementarities and estimate bidder markups using the estimated bidder valuations.

This test is important because, in the presence of the complementarities, the current allocation rules pose a hazard for bidders: the so-called "exposure problem". As a bidder must bid for each license before it knows whether it will win complementary licenses, it would be exposed to a risk of financial loss if the bidder bids more than its "stand-alone" value for a license. The stand-alone value of a license refers to the bidder's willingness-to-pay for a license in the case that it wins only that license. The exposure problem can hurt efficiency if efficient allocation requires the realization of complementarities. However, the exposure problem will be relevant only if complementarities exist.

Markups measure the bidders' pricing power exerted in the AWS-1 auction. In an auction game, large markups can arise because bidders strategically shade their bids below their valuations to exert oligopsony power in the presence of private information. Especially in a multi-object auction, the possibility that bidders can split objects among themselves at low prices could lead to large bidder markups.[4] Heterogeneity across bidders can be another source of market power. In the AWS-1 auction, the four largest winners, which include nationwide carriers such as T-Mobile and Verizon, won 71% of the licenses. To investigate whether these firms faced adequate competition, I estimate the distribution of private information on the values of the licenses and their markups.

Specifying the behavioral assumptions, I consider a bidder's decision to bid on a particular license at a given bid amount in an independent private value setting. The assumptions address the implications of the decision on the license's marginal contribution to the set of licenses the

[3] For example, as $1,122$ licenses were offered in the AWS-1 auction, the number of possible combinations of licenses, $2^{1,122} - 1$, exceeds the number of atoms in the universe!

[4] Ausubel and Cramton (1998) illustrate strategic demand reduction and Brusco and Lopomo (2002) consider tacit collusion in this context. Demand reduction arises when a bidder prefers to bid on fewer licenses than it desires in order to maintain low prices on the licenses upon which it is actually bidding.

bidder may win. The marginal contribution of a license to a set of licenses is defined as the bidder's willingness-to-pay to add the license to the set. This will be greater than the bidder's stand-alone value for the license if the set contains complementary licenses. The basis for the following assumptions is that the bidder will look ahead and compare the bidding price to the additional revenue from adding the license to the set of other licenses at each terminal node of the game.

The first behavioral assumption is as follows. If a bidder bid on a license at the given price, it implies that

BA1: the bidder expected to earn a positive surplus if it wins the license at the current price. In other words, the expected marginal contribution of the license to the set of licenses that the bidder may win was greater than the bidding price. BA1 allows a bidder to bid above the stand-alone value for a license. In this case, the bidder will not expect ex-ante to incur a loss if it wins the license at the bidding price, although it may do so if it fails to win the complementary licenses at the end of the auction.

The second behavioral assumption considers a pair of licenses (A, B) and asks why the bidder bid on license A, but not on license B, at the given prices of the two licenses. In this case, BA2 says that this bidding decision implies

BA2: the bidder preferred winning license A at the bidding price to winning license B at the price of B ex-ante.

In other words, the expected marginal surplus from winning license A at the bidding price was not smaller than that from winning license B at the price of license B. If a bidder, who was previously bidding on license A, stops bidding on A and starts bidding on license B, BA2 will arise twice: first, when the bidder bids on license A, but not on license B and second, when the bidder bids on license B, but not on license A. Examining these two decisions, the assumption attributes this switch to either (i) changes in the prices of the two licenses, or to (ii) changes in the bidder's perspective on the other licenses it may win at the end of the auction.

BA1 and BA2 constitute revealed preference inequalities consistent with rational behavior that should arise towards the end of an auction. While they have intuitive appeal, they are also weak restrictions. First, they do not require bidders to bid up to certain values on each license. Second, the assumptions allow bidders to engage in tacit collusion as long as a researcher can separate out the subset of licenses on which collusion occurred. Furthermore, the assumptions hold across existing leading models of bidding, including Brusco and Lopomo (2002), Milgrom (2000) and Zheng (2003).

While the assumptions specify implications of observed bidding on bidder valuations, they leave the data generating process unspecified. Therefore, they do not imply a unique distribution of bidder valuations given a distribution of bids. In other words, the assumptions identify only the set to which the true distribution belongs. This gives rise to a set estimator.

I assume that the complementarities among the licenses in a collection are captured by a function of bidder-collection specific characteristics. The observed bidder-collection specific variables, however, will not be enough to account for all the determinants of a bidder's willingness-to-pay for the collection. This creates the need to incorporate unobserved heterogeneity into the bidder's value for a bundle of licenses. I allow the bidder-license specific values to represent the unobserved heterogeneity. This flexible specification of unobserved heterogeneity requires an estimation strategy that can account for the fact that this structural error term, which affects bidding decisions, enters the revealed preference inequalities generated by the behavioral assumptions.

To estimate the coefficients that govern the magnitude of the complementarities, I employ an estimator developed by Pakes, Porter, Ho, and Ishii (2006) (hereafter PPHI). An advantage of this estimator is that it allows me to use a fixed effect approach to cancel out structural error terms. The behavioral assumptions make use of the panel data structure of bid data. It is a panel because bidders make repeated decisions to bid on the same licenses during the auction.

Once the structural errors are canceled out, the behavioral assumptions imply conditional moment restrictions that hold as an inequality. I use these conditional moment conditions to construct the unconditional moment inequalities that the set to which the true parameters belong should satisfy. After I recover the complementarities, I further estimate the distribution of the unobserved stand-alone values of each license.

In the application to the AWS-1 auction, the estimation results reject the hypothesis of no complementarities among AWS-1 licenses. This result is consistent with many empirical studies on synergies among spectrum licenses.[5]

For the estimation of the distribution of bidder-license specific values, I divide the bidders into two groups: global bidders, such as incumbent cellular phone carriers that operate on a large scale like T-Mobile and Verizon Wireless, and local bidders, such as rural telephone companies. The estimation results show that an increase in the Pop-MHz of a license, the population of the area covered by the license multiplied by the bandwidth of the license, increases a bidder's willingness-to-pay for the license for both bidder groups. This positive marginal effect of Pop-MHz on the stand-alone value of a license is more pronounced with the global bidders for the top 50 Metropolitan Statistical Area (MSA)s than for the other markets. The results

[5] Moreton and Spiller (1998) and Ausubel, Cramton, McAfee and McMillan (1997) find evidence for complementarities among licenses in the PCS auctions based on the reduced form regression of the winning bid for each license on a set of regressors. Bajari and Fox (2007) also report the existence of complementarities amongst the PCS C block licenses.

also show that the further away the market associated with a license is from a local bidder's location, the greater the negative effect will be on the mean value of the license.

I calculate the expected markups of the winning bidders using the estimated distribution of the stand-alone value of each license. The bidder markups are high and vary a lot across winners. While the markups of some winning bidders are barely positive, many have markups above 30%, implying that many winners paid only 70% of their values for their winning collections. This result is consistent with the following facts. First, the auction prices were below private transaction prices for similar spectrum licenses. Second, the number of bidders was not large compared to the number of licenses offered for sale. Despite there being over 160 bidders in the entire auction, competition was thin. Only a few global bidders and a small number of local bidders competed for any given license. The median for the local bidder group is 26% whereas the median for the global bidder group is 31%. This modest difference in the markups between the two groups suggests that the bidders were horizontally heterogenous because the complementarities were not large enough to create vertical heterogeneity.

This paper relates to the literature in the spirit of Haile and Tamer (2003). To estimate bidder valuation in single object English auctions, Haile and Tamer (2003) specify behavioral assumptions that are necessary conditions for a widely-accepted equilibrium, and that admit more general bidding behavior. Their incomplete model approach has been embraced by researchers who seek empirical structures to estimate bidder valuations in multi-unit auctions. (e.g. McAdams (2008), Chapman, McAdams and Paarsch (2006), Kastle (2008) and Hortacsu (2002)) This is because the theory of multi-object auctions lags far behind that of single object auctions, which guided the standard methods of estimating distributions of valuations (e.g. Guerre, Perrigne and Vuong (2000) and Donald and Paarsch (1993)). Especially for spectrum auctions, many economists agree that the interaction between the auction format and the complex nature behind demand for multiple heterogenous objects does not admit formal modeling in a general setting.[6]

As for the spectrum auction, the only attempt to estimate bidder valuations from auction data was made by Bajari and Fox (2007). They estimate bidder valuations based on the assumption that the total surplus of two bidders must not be increased by a pairwise exchange of licenses. The main differences between their work and mine lie in our respective goals and structures of unobserved heterogeneity. Bajari and Fox (2007) employ a matching game estimator that only uses final allocations. This requires a more stringent restriction on the structure

[6] Many theoretical models characterize equilibrium in a simultaneous ascending auction based on a model with two objects. Brusco and Lopomo (2002) considers a model with two objects and two bidders and extends it to one with two objects and N bidders. Zheng (2003) considers a model with two objects and three bidders. To remove the exposure problem, he assumes that bidders can withdraw their bids at no cost at the moment when jump bidding occurs.

of unobserved variables. The deterministic (observed) part of each bidder's value should be the major determinant of the final allocation. My estimation strategy allows the unobserved bidder-license specific values to be the major determinant by recognizing the panel structure of bid data. Their approach can be justified as they focus on ex-ante efficiency regarding the complementarities before bidders' stand-alone values for each license are drawn. My goal is to estimate the ex-post bidder markups in order to gauge the level of competition in a spectrum auction.

This paper contributes to the empirical industrial organization literature in the following ways. First, it contributes to the literature on spectrum auctions with a structural approach. In contrast to Bajari and Fox (2007), who recover the complementarities only, I recover the entire distribution of bidder valuations, more precisely a set to which the true distribution of bidder valuations belongs. This extended recovery enables me to address issues beyond the presence of complementarities such as bidder markups. This is the first paper that estimates bidder markups in an FCC spectrum auction. The large markup estimates in AWS-1 suggest that the distortionary effects of informational rents were large in the auction.

Second, this paper relates to the growing literature on partially identified models. The applied or theoretical econometric literature on estimation and inference of partially identified models has grown rapidly in recent years. I employ an estimator developed by PPHI and one similar to Manski and Tamer (2002)'s modified minimum distance estimator. This is the first paper that applies PPHI to auction estimation. Third, the approach proposed in this paper can be modified for discrete choice problems that cannot be analyzed using standard methods. The presence of complementarities among commodities gives rise to a combinatorial property in a consumer's choice problem. Under this environment, a consumer's choice set can be too large to admit a standard method. Fox (2007) and Rubinfeld, McCabe and Nevo (2006) consider such problems.

The remainder of this paper is organized as follows: Section 2 introduces the basic rules of FCC spectrum auctions with examples from the AWS-1 auction. In Section 3, I discuss the behavioral assumptions and present a simple model that provides a guide to solving a practical problem that arises from the assumptions. I describe the estimation approach and prove consistency of the proposed estimator in Section 4. Section 5 provides summary statistics of the AWS-1 auction in detail and discusses the parametrization choices for estimation. The estimation results are presented in Section 6. Section 7 concludes.

2 FCC Spectrum Auctions

In 1993, Congress passed a bill that gave the Federal Communications Commission the authority to use competitive bidding to allocate an initial license for the electromagnetic spectrum. Prior to this legislation, the Commission mainly relied on comparative hearings and lotteries to select a single licensee from a pool of applicants for a license.

Preparing an auction to allocate licenses to use a band of frequencies, the FCC specifies certain types of services for which the band should be used and some rules associated with how it can be used. For example, the frequency band that ranges from 1710 to 1755 MHz and from 2110 to 2155 MHz, labeled Advanced Wireless Service (AWS-1), was designed mainly for the third generation ("3G") mobile phone service.

Next, the FCC subdivides the band into smaller "blocks". Each block may be different in the bandwidth. For example, the FCC divided the band AWS-1 into five sub-blocks: A through F. Of these blocks, A, B and F were 20 MHz wide while the remaining blocks, C and D, were only 10 MHz wide. Blocks A, B and F, because they have twice the capacity, can potentially handle twice the number of simultaneous phone calls, or twice the data throughput.

Each block can be licensed to a single service provider as one nationwide license, as is the practice in many European countries. Alternatively, sub-blocks can be licensed to many service providers, each one serving a smaller geographical region, which is more typical in the US. In other words, a band of frequencies is not only divided by block, but also by geography. The geographic divisions for each block can be different. For example, for the AWS-1 band, the US and its territories were divided into 734 CMA(Cellular Market Area)s for block A while larger 12 REAG(Regional Economic Area Grouping)s for the D, E, and F blocks. Thus, allocation of licenses to use a band of frequencies amounts to allocation of multiple heterogenous licenses.

2.1 Auction Rules

2.1.1 Allocation rule and simultaneity

In a simultaneous ascending (SA) auction, a bidder submits bids for individual licenses simultaneously as the auction proceeds in rounds. The allocation rule is simple: the highest bidder of each individual license becomes the winner of the license. In contrast to this individual pricing rule, there may be interdependency among licenses in values through super-additivity in a bidder's valuation. The value of a bundle of geographically diverse licenses is greater than the sum of the stand-alone values of its component licenses. Particularly in the wireless service industry, there could be several factors that lead to super-additivity in a bidder's valuation: i) existence of a minimum investment requirement for installation of infrastructure necessary for service, ii) local synergy effects in reducing system management cost and iii) advantages in advertising as a provider of extended coverage.

This super-additivity in a bidder's valuation gives rise to complementarities among licenses. The existence of complementarities and the individual pricing rule complicates a bidder's strategy. For example, even if a bidder's value for a license depends on whether it wins another license, the bidder must bid on individual licenses without knowing whether it will acquire the other. Therefore, a bidder who has pursued a large package, but was unable to win, may be left with a partial package whose total price cannot be justified without those complementary licenses.

This exposure problem stands at the center of complexity of the SA auction complicating a bidder's strategy, and hence the characterization of equilibrium in a general setting. Faced with the exposure problem, if a bidder stops bidding for a license before the price of the license reaches its marginal contribution to the package of licenses it wins, the resulting allocation is not only ex-post sub-optimal for the bidder but also potentially inefficient.[7]

To mitigate the exposure problem and allow bidders to adjust bidding policies and assemble packages as the auction progresses, bidding for all licenses opens and closes at the same time. This rule is called a simultaneous stopping rule. A FCC auction closes after the first round in which no bidder submits a new bid on any license.

2.1.2 Activity Rules

By implementing the simultaneous stopping rule, the FCC has adopted an activity rule to control the pace of an auction. The activity rule specifies a minimum level of activity for a bidder in order to maintain its *eligibility*, which determines ability to bid on licenses in subsequent rounds. A bidder's *activity* in a particular round is measured as the sum of the *size* of the licenses on which it is *active* in the round: a bidder is said to be active on a license if it is either the standing high bidder from the previous round or placing a new bid in the round. The size of each license is quantified as the number of *bidding units*. Bidding units for a license are typically determined based on the license's Pop-MHz, the population in the area associated with the license multiplied by the bandwidth size. For example, if a license with a bandwidth of 20 MHz covers an area with a population of 2 million, the license is characterized by 400

[7]Whether the exposure problem is a major issue is out of this paper's focus and scope. Klemperer (2000) does not include the exposure problem in his list of practical considerations in designing a multi-unit auction. Englmaier et al (2006) report that in their experiments, an auction design that imposes the exposure problem on bidders yields a higher (similar) revenue, depending on whether bidders are experienced (unexperienced), but yields a lower allocative efficiency compared to the second price sealed bid auction without the exposure problem.

million Pop-MHz.[8]

A bidder's *eligibility* defines the maximum level of activity the bidder can hold during the auction. Therefore, bidders need to ensure that they have enough eligibility to cover all the licenses they wish to purchase. A bidder can purchase as much initial eligibility as it wishes by making an upfront payment before the auction begins. A typical spectrum auction specifies the activity requirement as a percentage of the bidder's eligibility. For example, the AWS-1 auction required bidders to be active above at least 80% of their eligibility for round 1 to round 30, and then above 95% from round 31 on. Whereas a bidder's eligibility cannot be increased after the auction starts, it is reduced if it violates the activity requirement. For example, in AWS-1, the following formula is used to calculate a bidder's eligibility in the subsequent round based on its activity in the current round.

$$\textit{eligibility in round } r+1 = \textit{min}(\textit{eligibility in round } r, \frac{\textit{activity in round } r}{\textit{requirement percentage in } r})$$

Hence, if a bidder with a eligibility of 20 million bidding units in round 25 holds an activity level of 14 million bidding units, its eligibility will be reduced to 17.5(= 14/0.8) million bidding units. The activity requirement typically rises as the auction progresses.

2.1.3 Information Disclosure Rule and Other Rules

Each round consists of two periods: the placement of bids or withdrawals and the announcement of the standing high bids and other information. For auctions conducted since 1994, the FCC has adopted the full information disclosure rule.[9] Under this disclosure rule, the FCC posts all the relevant information of the on-going auction after every round. It includes the bids placed by each bidder on each license, the identity of the bidder, and the change in each bidder's eligibility.[10]

Bid withdrawals are allowed, but subject to the bid withdrawal payments specified by the FCC.[11] In the AWS-1 auction, the number of rounds in which a bidder can withdraw its

[8] The license's Pop-MHz is the single most important component that determines value of the license since greater Pop-MHz implies a larger pool of potential service subscribers and capacity.

[9] An exception was made when the FCC used anonymous bidding for a recent auction, labeled 700 MHz, in 2008.

[10] Full information disclosure has created some concern over anti-competitive bidding. The FCC announced it would conduct the AWS-1 auction under limited information disclosure unless the gauge of the likely level of competition is equal to or greater than 3. This gauge was measured as the ratio of the sum of all the bidders' initial eligibility, measured in bidding units and subject to the cap, to the sum of bidding units of all the licenses offered for sale. The FCC conducted the AWS-1 auction under the usual full information procedures as the ratio turned out 3.04.

[11] For the case of a single withdrawal on a license, the bidder must pay the difference between its withdrawn bid and the subsequent winning bid either in the auction where the withdrawal is made or in subsequent auctions.

standing high bids was limited to two rounds per bidder. Withdrawals were allowed to reduce the risk associated with efforts to secure various licenses in combination. Only a small number of withdrawals, however, were placed in several major auctions including AWS-1. This indicates that the risk to which bidders were exposed was not severe.

As another device for controlling the pace of an auction, the FCC specifies the *minimum acceptable bid* for each license and round. In order for a new bid placed on a license to be counted towards the bidder's activity, it must exceed the minimum acceptable bid of the license for that round. The minimum acceptable bid is determined by applying a minimum bid increment to the standing high bid from the previous round. The AWS-1 auction began with positive minimum opening bids. In AWS-1, bidders were allowed to place a bid that strictly exceeded the minimum acceptable bid, i.e., a jump bid, by choosing from the given eight additional acceptable bids. These additional acceptable bids were determined by multiplying and then rounding the minimum acceptable bid by successively larger numbers such as 1.1, 1.2, and so on. Jump bids did not frequently occur in major auctions including AWS-1. This implies the strategic effects of jump bidding, as addressed in Zheng (2003) and Avery (1998), were unlikely to be important.

3 Behavioral Assumptions

The goal of this section is to specify behavioral assumptions that are reasonable, and that allow for the presence of complementarities. I consider what a bidder's decision to bid on an individual license reveals about its value for the license. More specifically, I relate this decision to the marginal value of the license at each terminal node of the auction game, which corresponds to a different final price vector. As the bidder's surplus-maximizing collection of licenses should be different at each final price vector, the value from adding the license will also be different. In this environment, I assume that the decision implies the ex-ante expected marginal surplus from winning the license at the bidding price is positive. I further assume that the decision, compared to an alternative decision to bid on a license that the bidder did not bid on, is expected-surplus-enhancing.

The main premise behind the validity of these assumptions is that an alternative decision does not change the distribution of the final price vector. I assume that an econometrician knows a subset of licenses for which this premise is satisfied. I discuss the identification of this set for the AWS-1 auction in detail in section 3.3 based on implications from the model in section 3.2.

3.1 The Assumptions

Let $\mathcal{T}$ denote the set of licenses for sale and $V_i(S)$ bidder i's dollar valuation for collection S. $\pi_i(S) = V_i(S) - \sum_{l \in S} p_l$ denotes i's surplus from winning collection $S \subseteq \mathcal{T}$ given the price of license l as p_l. Bidder i's stand-alone value of license l is denoted by $v_{il} = V_i(\{l\})$, which reflects bidder i's profit from providing the wireless service in the area associated with license l only. I assume that bidder i's valuation for collection S can be decomposed into two parts: the sum of the stand-alone values of the component licenses $\sum_{l \in S} v_{il}$ and the term that captures the complementarities among the licenses in the collection k_{iS}. That is,

$$V_i(S) = \sum_{l \in S} v_{il} + k_{iS}$$

I assume that spectrum licenses, or a bundle of spectrum licenses, have private values so that other bidders' signals or values do not affect the bidder's value for the license or a bundle that includes the license.[12] I also assume that there is no information asymmetry regarding the magnitude of the complementarities $(k_{iS})_{S \subseteq \mathcal{T}}$. The only private information is each bidder's stand-alone value for each license $(v_{il})_{l \in \mathcal{T}}$. This assumption is not too restrictive because the main sources for the complementarities are likely to be advantages in system management and production costs due to economies of scale in the wireless service industry.

Marginal contribution of license l at collection S is defined as:

$$\Delta V_i(l, S) = V_i(S \cup \{l\}) - V_i(S \backslash \{l\})$$

so that it corresponds to bidder i's willingness-to-pay for additional license l if i already owns collection $S \backslash \{l\}$. By this definition, license l's marginal contribution to any collection for bidder i includes i's stand-alone value for license l. If collection S contains a license complementary to l, the marginal contribution of l to S will be greater than l's stand-alone value. If bidder i wishes to purchase only one of license l and license l', the marginal contribution of l to a collection that includes l' will be zero. I assume the marginal contribution of a license to any collection is non-negative.

p_{ilt} denotes the minimum price bidder i must pay to win license l after the auction has proceeded to round t. Therefore, p_{ilt} equals the provisionally winning bid if i is the current standing high bidder on l in round t and otherwise the minimum acceptable bid. By definition,

[12] It is possible that some bidders' information is useful to assess values of a collection of licenses to the other bidders. In our empirical context, however, this effect would be minimal because demand for wireless services of each market can be easily analyzed without much variation and each bidder may have a very different business plan, which likely plays a more important role in determining profitability of each license.

p_{ilt} is common across bidders except one bidder who is the standing high bidder on license l in the beginning of round t.[13] Let $P_{iSt} = \sum_{l \in S} p_{ilt}$ denote the minimum price bidder i must pay to win collection S in round t. For simplicity of notation, in the following, I omit subscript i from p_{ilt} unless I need to denote bidder i is the standing high bidder on l. I assume that a bidder takes the bid amount for license l given as p_{lt}. This simplification is justified as most of the bids in FCC spectrum auctions, including AWS-1, were exactly the minimum acceptable bids.[14]

T denotes the final round and H_{iT} the collection that maximizes bidder i's surplus given final prices of licenses $(p_{lT})_{l \in \mathcal{T}}$. Hence, the probability distribution of H_{iT} should be determined by bidder i's (equilibrium) belief on the final prices of licenses. J_{it} denotes bidder i's information set in the beginning of round t. This information set includes H_{it}, the set of licenses bidder i is provisionally winning in the beginning of round t. B_{it} denotes the set of licenses for which bidder i bid in round t.

The following behavioral assumptions relate a bidder's behavior of bidding on a license at the given bid amount to the expected marginal contribution of the license.

Behavioral Assumptions *There is round r_0 and a set of licenses $\mathcal{R} \subseteq \mathcal{T}$ such that if $l \in B_{it}$, $l' \notin B_{it}$ and $l,\ l' \in \mathcal{R}$ where $t \geq r_0$,* **BA1** *and* **BA2** *hold.*

BA1 $E(\Delta V_i(l, H_{iT}) |\, J_{it},\ p_{lT} = p_{lt}) \geq p_{lt}$

If bidder i bids on license l, the expected marginal contribution of the license to the surplus maximizing final collection is greater than the bidding price. The expectation is conditional on the current information set J_{it} and on the event that the final price of the license is equal to the current price $p_{lT} = p_{ilt}$.

BA1 can be viewed as a simultaneous ascending auction analogy of Assumption 1 in Haile and Tamer (2003). If a bidder desires a single license, BA1 corresponds to a necessary condition for the weakly dominant strategy in a single item button auction of Milgrom and Weber (1982). In this case, BA1 simply says that if a bidder bids on a license, the value of the license is greater than the bid. For a bidder who desires multiple licenses, BA1 allows the bidder to bid above its stand-alone value for the license. In this case, the bidder will not expect ex-ante to incur a loss if it wins the license at the bidding price although it may incur a loss if it fails to win the complementary licenses at the end of the auction. The expectation arises because even when the final price of license l is given as the current price p_{lt}, the value of the license depends on the set of other licenses the bidder will win.

[13] Milgrom (2000) terms p_{ilt} bidder i's *personalized price* for license l in round t.

[14] Bajari and Yeo (2008) provide descriptive statistics on this matter. I also assume that a bidder does not bid on licenses on which it is the high standing bidder. Bajari and Yeo (2008) report that the cases in which the provisionally winning bidder increases its own bid in the auction are infrequent.

The condition that $p_{lT} = p_{lt}$ captures the fact that the bidder only needs to account for the cases where it is stuck with license l at the end of the auction. If bidder i is outbid on the license in the future, the price of the license it faces will increase above p_{lt}. At that point, bidder i can choose not to bid on license l.[15]

I further assume the following:

BA2 $E(\Delta V_i(l, H_{iT})|J_{it}, (p_{lT}, p_{l'T}) = (p_{lt},\ p_{l't})) - p_{lt}$
$$\geq E(\Delta V_i(l', H_{iT})|J_{it}, (p_{lT}, p_{l'T}) = (p_{lt},\ p_{l't})) - p_{l't}$$

If bidder i bids on license l, but not on license l', the expected increase in the surplus from winning license l at the bidding price is not smaller than that from winning license l' at the current price of l'. If a bidder desires a single license, BA1 implies BA2. The condition that $(p_{lT}, p_{l'T}) = (p_{ilt},\ p_{il't})$ captures the fact that the bidder only needs to account for the cases where the final prices of the two licenses are the current prices. In those cases, bidder i would be stuck with the license it chose to bid on, while missing the chance to win the license it does not bid on at the possible minimum price.

BA2 tries to capture the implications of the occurrence of "substitutions" between the licenses a bidder bids on during the auction. If a bidder, who was previously bidding on license l, stops bidding on l and starts bidding on license l', BA2 will arise twice: first, when the bidder bids on license l, but not on license l' and second, when the bidder bids on license l', but not on license l. Examining these two inequalities, the assumption attributes this switch to either (i) the changes in the prices of the two licenses, or to (ii) the changes in the bidder's perspective on the other licenses it may win at the end of the auction.

BA1 and BA2 require that the bidder's perspective on the prices of licenses other than the one being considered do not change whether it bids on the license or not. For example, BA1 compares the expected surplus when the bidder wins license l at price p_{lt} with the expected surplus when l is taken out of the final collection. That is, BA1 can be rewritten as $E(V_i(H_{iT} \cup \{l\}) - \sum_{j \in H_{iT} \cup \{l\}} p_{jT} \Big| J_{it}, p_{lT} = p_{lt}) \geq E(V_i(H_{iT} \backslash \{l\}) - \sum_{j \in H_{iT} \backslash \{l\}} p_{jT} \Big| J_{it}, p_{lT} = p_{lt})$. Comparing the left and right hand side of the inequality shows that the distribution of the final prices are fixed regardless of whether the bidder bids on l or not. The same argument applies to BA2. BA2 requires that the bidder believes bidding on license l or license l' won't change the final prices of other licenses. Therefore, BA2 may not hold if, for example, bidder i did not bid on license l' because doing so could trigger a different equilibrium or raise the final prices of other licenses.

[15] I implicitly assume that a bidder is not allowed to withdraw its standing high bid. If bid withdrawals were allowed, a bidder must consider the possibility of withdrawing its bid when placing a bid. Therefore, it would consider the final price of the license it bids on beyond the current price.

This is why BA1 and BA2 are assumed to hold only for $l \in \mathcal{R}$ and $l' \in \mathcal{R}$. That is, I assume that bidders' bidding decisions on licenses in $\mathcal{R} \subseteq \mathcal{T}$ do not trigger intensive competition that would otherwise not arise. To identify $\mathcal{R} \subseteq \mathcal{T}$ for the AWS-1 auction, I turn to implications of the simple model in the following section.

The reason why BA1 and BA2 are assumed to hold only for rounds later than round r_0 is as follows. BA1 and BA2 are based on the premise that bidder i believes the probability of winning a license at the license's current price is strictly positive. This may not be true in early rounds of the auction because the prices of the licenses are relatively low. If a bidder knew the bidding prices of some licenses would not be the final prices, the bidder could bid on those licenses without considering the implication of the decision on its expected surplus from winning the licenses. An example of this kind of bidding behavior will occur if bidders delay bidding for licenses they truly desire and bid on many licenses intermittently in order to meet the activity rule.[16] Such behavior is suppressed more effectively in late rounds by relatively strict activity requirements. In later rounds, there is a high probability that the bidder wins the license upon which it bids at the bidding price. Therefore, the risk of bidding on unwanted licenses and the expected opportunity cost of not bidding on profitable licenses will be greater in later rounds.

These assumptions will be not be informative as to the magnitude of the complementarities if the bid amount for each license does not exceed each license's stand-alone value. Therefore, restricting attention to only rounds after round r_0 is not likely to result in a significant loss of information.[17]

While BA1 and BA2 have intuitive appeal, they also have the advantage of being weak restrictions. First, they do not require bidders to bid up to certain values on each license. Second, the assumptions allow bidders to engage in tacit collusion as long as a researcher can separate out the set of licenses $\mathcal{T} \setminus \mathcal{R}$ on which collusion might have occurred.

[16] This behavior is called parking. Parking behavior describes a strategy in which a bidder, who wishes to hide its licenses of interest until later rounds, bids for licenses it is not interested in purchasing only to satisfy the activity requirement in early rounds. Salant (1997) documented this behavior based on his experience as a consultant for GTE in the PCS A&B block auction. He argues that the GTE bidding team has engaged in the parking strategy to stop other bidders observing the identities of GTE's licenses of interest and then raising bids on those licenses to take advantage of GTE's liquidity constraint. Bajari and Yeo (2008) reports that there is no evidence for the parking behavior in the auctions with relatively high opening bid requirements including AWS-1.

[17] Also, if a bidder is risk averse, it will be less reluctant to bid above the stand-alone values for licenses in later rounds. The exposure problem is less severe in later rounds as bidders have a better sense of the final allocation. In this case, a bidder's willingness-to-pay for a license depends not only the license's expected marginal contribution but also the variance of the distribution of the final collection. I assume that a bidder's (Bernoulli) utility function is linear so that its willingness-to-pay for a license, given the distribution of the collection it may win, depends only on the expected marginal contribution of the license. Therefore, I do not explicitly differentiate the willingness-to-pay from the expected value of the additional revenue generated by a license.

If licenses are *mutually substitutes* and bidders *bid straightforwardly* as in Milgrom (2000), the behavioral assumptions will hold. If licenses are independent in their values and bidders bid straightforwardly, these assumptions will hold trivially. In this case, each bidder will place a bid on a license as long as the bid amount is not greater the license's stand-alone value. The assumptions also hold in the behavior described in Zheng (2003)'s Lemma 1.

They also hold in a collusive equilibrium with and without complementarities as described in Brusco and Lopomo (2002). However, in the equilibria described in Brusco and Lopomo (2002), bidders do not bid on a license at a bidding price that exceeds the stand-alone value of the license. In this case, the assumptions will not be informative in detecting the presence of the complementarities, let alone the magnitude. The model in the following subsection presents (collusive) equilibria where bidders bid on a license at prices larger than their stand-alone values for the license.

3.2 An illustrative model

In this section, I present a simple model with two bidders and three licenses. The model serves two purposes. First, it provides a concrete example of bidding behavior in the simultaneous ascending auction environment. In particular, I characterize equilibria where bidders bid on a license at prices that exceed their stand-alone values. This occurs on only one license that remains after the other two licenses have been allocated. Second, the model shows that when there are more licenses than bidders, both collusion and competition occurs. The collusion stage precedes the competition stage. I use this implication to determine the set of licenses $\mathcal{R}$ in AWS-1 for which I will apply the revealed preference inequalities, BA1 and BA2. I discuss this issue in greater detail in section 3.3.

Brusco and Lopomo (2002) characterize a collusive equilibrium in a simultaneous ascending auction with two bidders and two objects and extend it to the case of N bidders and two objects. In the latter case, bidders bid only on their most preferred license and then raise their bids competitively until only two bidders remain. At that point, the remaining bidders split the two objects between themselves unless their most preferred objects are the same. In their collusive equilibria, no bidder bids more than the stand-alone value of a license at any point during the auction. The model in this section provides a case where bidders bid on a license at prices larger than the stand-alone values.

The key difference between the model in this section and Brusco and Lopomo (2002)'s is that the number of licenses exceeds the number of bidders, which is typically the case in a spectrum auction. In this case, bidders cannot split the licenses equally amongst each other. In the equilibria, each bidder will acquire one license at a low price via collusion and then "compete" for the remaining license. Each bidder will be willing to bid up to the marginal contribution of the remaining license to its final collection which contains the license it has acquired via collusion.

Let $i = 1, 2$ index bidders, and $\mathcal{T} = \{a, b, c\}$ denote the set of licenses offered for sale. Furthermore, I assume that the magnitude of complementarities is non-negative and common across bidders, that is $k_{i,S} = k_S$ for all $i = 1, 2$ while the stand-alone values $\{v_{il}\}_{l \in \{a,b,c\}}$ are privately known to bidder i. The stand-alone values are independently drawn from the same probability distribution F with support $[0, 1]$. I further assume $k_{\{a,b\}} = k_{\{a,c\}} = k_{\{b,c\}} = k$ to simplify the analysis. This assumption ensures that the licenses are symmetric in their contributions to the complementarities. I also assume that $k_{\{a,b,c\}} = K > k + 1$.

The auction proceeds in rounds in which each bidder can raise the standing high bid by at least a minimum increment. The standing high bid for a license in round r is defined as the highest bid placed in round $r - 1$. The auction ends after the first round in which no bidders place a new bid on any license. A minimum increment is assumed to be negligible, or close to zero. Furthermore, let a bid of $-\infty$ denote "no bid".

The next proposition establishes the existence of a competitive equilibrium which will serve as a threat to sustain a collusive equilibrium.

Proposition 0 *There exists a perfect Bayesian equilibrium in which the three objects are allocated to the bidder with the highest $v_i = V_i(\mathcal{T})$ at a price equal to the second highest valuation.*

Following Brusco and Lopomo (2002)'s intuition, if bidders compete on all licenses, then the auction cannot end with a bidder winning only two licenses. The reason is that, if a bidder has won two licenses, then the value of the other license is at least $K - k > 1$. Since this is more than the largest value that a winner of a single license is willing to pay, a winner of two licenses will pursue the third license. Therefore, both bidders behave as if they were bidding for a single object, the collection of three licenses.[18] The next proposition establishes a collusive equilibrium which yields a higher expected surplus for both bidders. In this collusive equilibrium, bidders use the equilibrium in Proposition 0 as a threat to sustain collusion.

Let $b_i(\mathcal{T})$ denote the bids that bidder i submits for each license and L_i the license that has the highest stand-alone value to bidder i.

[18] This is not true if there are moderate complementarities, i.e., $K < k + 1$. It seems well understood among auction theorists that it is difficult to characterize non-collusive equilibrium when complementarities exist. Brusco and Lopomo (2002) admits that it is difficult to characterize the competitive Bayes Nash equilibria even for the case with two objects and private, moderate complementarities. See Sherstyuk (2002, 2003) for existence and implementation of the competitive equilibrium in a SA auction with complete information.

Proposition 1 *Assume that* $E(v_l) \geq \dfrac{1+\gamma}{3}$ *where* $\gamma = \int_0^1 F(x)^3 dx$. *The following strategy with some consistent belief system forms a symmetric perfect Bayesian equilibrium:*

- i *opens with* $b_i(L_i, \mathcal{T}\backslash\{L_i\}) = (0, -\infty, -\infty)$.
- *If* $L_1 \neq L_2$, *each bidder raises its bid on the third object* l *such that* $l \notin \{L_1, L_2\}$ *until either i) the price reaches the value* $\beta_i = \Delta V_i(l, \{L_i\})$ *or ii) the opponent stops. In either case, they do not bid for the next round.*
- *If* $L_1 = L_2$, *bidders play the equilibrium in Proposition 0*
- *If, at any stage, a bidder makes a bid that cannot be observed under the strategy described above, the bidders play the equilibrium in Proposition 0.*

Proof See Appendix A. ■

The behavior implied by the equilibrium of Proposition 1 can be described as follows. Each bidder places a bid of the minimum acceptable bid (zero) on its "favorite" license that has the highest stand-alone value. If the bidders' favorite licenses are different, they stop raising their bids on those licenses so that each bidder becomes the standing high bidder of its favorite license. In the following rounds, each bidder raises its bids on the remaining license until either the minimum acceptable bid exceeds the marginal contribution of the license to its winning collection, or the opponent stops. For example, if bidder 1 opens with a bid on a and 2 on b, bidder 1 will bid until the minimum acceptable bid on c exceeds the marginal contribution of license c to license a or bidder 2 stops bidding. This strategy is optimal since the exposure problem is resolved because each bidder knows exactly what license it will win at the end of the auction, except the license on which it is bidding. Therefore, the sub-game from the second round is reduced to a single item English auction.[19] If their favorite licenses are the same, bidders revert to the equilibrium strategy in Proposition 0.

The following proposition establishes another equilibrium where bidders do not trigger the competitive equilibrium even when they bid on the same license in the first round.

Proposition 2 *Under Condition A, the following strategy, together with some consistent belief system, forms a symmetric perfect Bayesian equilibrium:*

- i *opens with* $b_i(L_i, \mathcal{T}\backslash\{L_i\}) = (0, -\infty, -\infty)$.

[19] It is interesting that the sufficient condition for a collusive equilibrium with three objects is less restrictive than the similar equilibrium with two objects described in Proposition 1 of Brusco and Lopomo (2002). The existence of an opportunity to compete over the third object facilitates collusion more easily.

- *If $L_1 \neq L_2$, each bidder raises its bid on the third object l such that $l \notin \{L_1, L_2\}$ until either i) the price reaches the value $\beta_i = \Delta V_i(l, \{L_i\})$, or ii) the opponent stops. In either case, they do not bid for the next round.*

- *If $L_1 = L_2$, bidders raise their bid on the object l such that $l = L_1 = L_2$ until either i) the price reaches the value $\alpha_i = \min\{v_{iL_i} - v_{il} | \, l \neq L_i\}$, or ii) the opponent stops. In case i), the bidder bids on l_i such that $v_{il i} = v_{iL_i} - \alpha_i$ and in case ii) the bidder stops bidding at the following round. As soon as both bidders become the standing high bidders of different licenses, denoted by (l_1, l_2), bidders raise their bid on the third object l such that $l \notin \{l_1, l_2\}$ until either i) the price reaches the value $\beta_i = \Delta V_i(l, \{L_i\})$, or ii) the opponent stops. In either case, bidders do not bid for the next round.*

- *If, at any stage, a bidder makes a bid that cannot be observed under the strategy described above, the bidders play the equilibrium in Proposition 0.*

Proof See Appendix A. ■

In the equilibrium in Proposition 2, both bidders signal their valuations to derive as high a surplus as possible from acquiring a single license as in Proposition 1. Proposition 2 differs from Proposition 1 in that if both bidders prefer the same license at the opening price, they raise their bids on that license until one bidder becomes *indifferent* between the license and its second-favorite license that has the second highest stand-alone value. After both bidders become the standing high bidder of one license, they raise their bids on the remaining license as in Proposition 1.[20]

Complementarities do not play a role because they are assumed to be common across bidders and are hence competed away. The assumption that $k_{\{a,b\}} = k_{\{a,c\}} = k_{\{b,c\}}$ ensures that licenses are symmetric in their contributions to the complementarities so that bidders only consider the stand-alone value for each license when choosing the license they may acquire at the price of zero.[21]

[20] The current activity rule may not be able to prevent the collusive equilibrium from arising. For example, if the activity rule requires that bidders remain active above at least 50% of their eligibility and the bidding units are the same across the licenses, the bidders can meet the rule by starting with the eligibility level only enough to buy two licenses. In this case, an initial purchase of eligibility is a part of a bidder's strategy. If a bidder's initial eligibility is large enough to cover all three licenses, bidders would revert to the competitive equilibrium.

[21] To see how a different assumption affects an equilibrium strategy and complicates the analysis, consider a case with $k_{\{a,b\}} > 0$ and $k_{\{a,c\}} = k_{\{b,c\}} = 0$. Suppose that bidder 1's stand-alone values for licenses are $(\alpha - \epsilon, \alpha - \epsilon, \alpha)$. Bidder 1 has an incentive to deviate from the first round equilibrium behavior $b_1(a, b, c)$ $= (-\infty, -\infty, 0)$ to $b_1(a, b, c) = (0, -\infty, -\infty)$. The incentive is that if bidder 2 opens with a bid on c, bidder 1 could outbid bidder 2 on b because its willingness to pay for b, given it is the standing high bidder of a, is likely to exceed bidder 2's stand-alone value for b and hence enjoy complementarities. In this case, the license for which each bidder submits a bid in the first round should be determined simultaneously in equilibrium.

In general, licenses are different as the population and capacity associated with each license vary. If some licenses are more important in determining the boundaries of collections that bidders will pursue during the auction, bidders will have an incentive to determine the winners of those licenses in early rounds. By doing so, bidders can facilitate tacit collusion similar to the one in this section. Also, they can mitigate the exposure problem. I present suggestive evidence on this point using data from AWS-1 in the following section.

3.3 Discussion

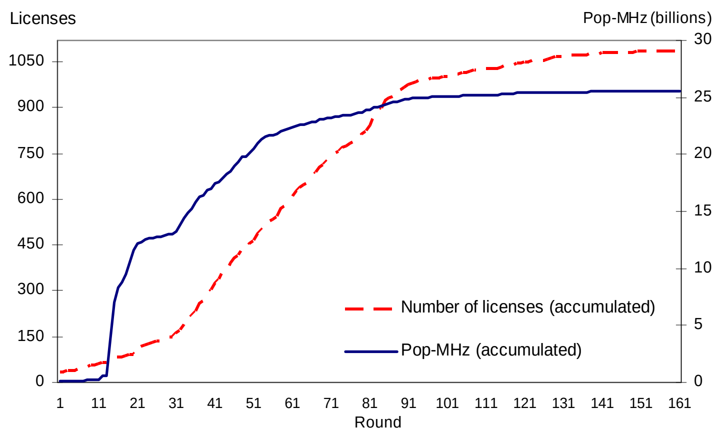
One of the behavioral implications from the equilibrium in Proposition 2 is that a bidder may stop bidding on its favorite license even though the price is below its willingness-to-pay. This happens because the bidder lets the opponent win the license in the collusive equilibrium. As the price rises on the remaining license that the bidders are competing over, the expected surplus from the bidder's favorite license potentially becomes larger. However, the bidder would not bid on the license it let the opponent win so as not to trigger the competitive equilibrium. Therefore, applying BA2 to this pair of licenses will generate a false inequality. To prevent this case from arising, I assume the researcher can identify licenses for which collusion occurs and exclude them from $\mathcal{R}$.

To determine $\mathcal{R}$ for AWS-1, I turn to the implications of the model. The model shows that the collusion stage precedes the competition stage. This is because competition occurs on the residual licenses whose winners were not determined via collusion. I extend this implication to AWS-1.

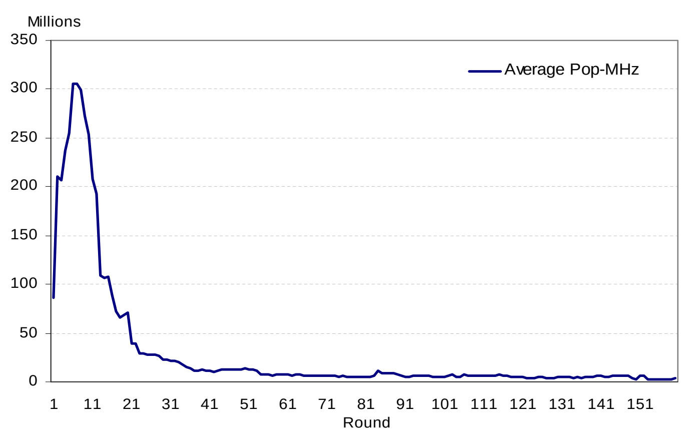
In Figure 1 (a), the straight line charts the cumulative number of licenses that received winning bids up until each round. The dashed line shows this information in terms of the cumulative amount of Pop-MHz. The winning bids for almost half of the licenses for sale have been placed before round 60. It also shows that the winners for licenses with large Pop-MHz tend to be determined in earlier rounds compared to licenses with smaller Pop-MHz. Figure 1 (b) presents this phenomenon in a different way. According to the figure, in early rounds, new bids are more concentrated on licenses with large Pop-MHz so that the winners of those licenses are determined in relatively earlier rounds.

Licenses characterized by large Pop-MHz include licenses that cover a broad area and a small area with a high population density such as top metropolitan statistical areas. Intuitively, licenses with large Pop-MHz are likely to be more important in determining a profitable collection of licenses a bidder would like to pursue during the auction. Therefore, the incentive to raise prices on those licenses to determine who wins them will be shared among bidders.[22] I

[22]It is not clear whether bidders have an incentive to determine the winners of large licenses solely due to the



(a) Place of winning bids



(b) Average Pop-MHz of new bids

Figure 1: Bidding pattern in AWS-1 over rounds

take this as evidence that if collusion had occurred in AWS-1, it would be on "large" licenses in early rounds. Based on this observation, I restrict the set of licenses $\mathcal{R}$ for which I apply BA1 and BA2 to be the set of "small" licenses that have received a bid from the bidder in a late round. In AWS-1, small licenses refer to CMA licenses that divided the US into 734 small areas.

This is also more consistent with the substitution patterns BA2 tries to capture. According to BA2, a bidder switches to a new license only if it provides a greater expected marginal surplus. This happens because prices are changing or because the bidder updates its belief about other licenses it will win.

Suppose that a profitable collection for a bidder who desires multiple licenses changes as the bidder's perceived likelihood to win large licenses changes. Under this circumstance, the bidder may wish to delay bidding on small licenses until it is certain about large licenses it will win in order to mitigate the exposure problem. However, the activity rule forces the bidder to bid on small licenses as well while raising its bids on large licenses. Therefore, as the profitability of collections changes, the small licenses the bidder bids on will change as well.

Another implicit implication from the simple model and many existing models is that once a bidder starts bidding for a license, it bids on the license continuously up to some price level. Although there are many bidders who exhibit this bidding pattern in AWS-1, this is not generally true. It is observed that some bidders bid on many licenses intermittently throughout an auction moving back and forth from one license to another. One may be able to incorporate this bidding pattern by introducing liquidity constraints. If a bidder is liquidity constrained so that it can only add one more license to the set of licenses it is likely to win, it will have to choose one among all the licenses that have positive expected marginal surplus at current prices. Hence, it may bid on several licenses intermittently. The behavioral assumptions allow for such bidding behavior. However, this kind of substitution, induced only by changes in the prices of licenses that the bidder is bidding on, is not likely to provide information on the magnitude of the complementarities. Therefore, when I apply BA2 to a pair of licenses, (l, l'), I require two conditions: (i) bidder i never placed a bid on license l while bidding on license l' and (ii) bidder i never placed a bid on l' while bidding on l.

It is noteworthy that the assumptions would not say anything about a bidder's valuation for a license for which it is the standing high bidder, unless it places a new bid on the license in a round later than r_0. The reason is that changes in the bidder's information and hence its

activity rule. One supporting story would be that a bidder can meet the activity requirement more easily if it is the standing high bidder on a large license if it wants to hide its intention until later rounds. One story against it would be that the bidder would not be able to take advantage of a situation in which competition over some licenses turns out to be softer than expected if it lacks free eligibility because it is stuck with large licenses.

perspective on the distribution of H_{iT} could drive the expected marginal contribution of the license below the bidding price.[23]

4 Estimation Procedures

In this section, I describe the estimation procedures. I maintain the assumption that the value of a bundle of licenses is the sum of two terms: the sum of the stand-alone values of the component licenses and the complementarities. I assume that the stand-alone value of each license is private information and is not observed by an econometrician. I estimate the complementarity term and the distribution of the stand-alone values of a license separately 1through two stages.

In the first stage, I recover the complementarities by assuming the magnitude of the complementarities is a linear function of bidder-collection specific characteristics. The behavioral assumptions, together with this parametrization, lead to conditional moment restrictions that hold as an inequality. These are used to construct unconditional moment inequalities. To account for the presence of unobserved stand-alone values of licenses, which affect the bidding decision, I use a standard panel-data technique, differences-in-differences. An estimator developed by PPHI is applied.

In the second stage, the distribution of unobserved stand-alone values of each license is estimated. I modify the behavioral assumptions to capture the fact that the set of licenses that a bidder is either winning or is placing a bid on in a late round is likely to be the bidder's final collection. The modified behavioral assumptions lead to inequalities that generate upper and lower bounds of the stand-alone value of a license. The distribution of the stand-alone value is bounded by the distributions of these upper and lower bounds. I estimate the distribution of the stand-alone value of each license by parameterizing the distribution as a normal distribution with an unknown mean and variance that are functions of bidder-license specific variables. This parametrization is necessary for the following reasons. First, observations for each individual license are not enough to estimate the bounding distributions non-parametrically. Second, I need to extrapolate the distribution of the stand-alone value of a large license, which is not used in the estimation, from the distribution of the stand-alone value of small licenses.

[23]Imagine that a bidder who is the standing high bidder for license, say l_1, bids on license l_2 more than its stand-alone value because license l_2 is complementary to license l_1 and it expected to win $\{l_1\}$ with a high probability. Suppose an opponent places a new on l_1 at the price that bidder i does not want to include l_1 in its final collection any more and hence bidder i would be left with license l_2 only. Since the expected marginal contribution of l_2 at the point is bidder i's stand-alone value for l_2, the inequality in BA1 would fail.

4.1 First stage: PPHI (2006)

I begin this section by stating the assumption that parameterizes the value of a collection of licenses to a bidder.

Assumption 1 (Parametrization) $\pi_{it}(S) = x_{iS}\beta_0 + \sum_{l\in S} v_{il} - \sum_{l\in S} p_l$ *where* v_{il} *is private information.*

I parameterize the magnitude of complementarities among licenses in a collection S as a linear function of a vector of bidder-collection specific characteristics x_{iS} with dimension $1 \times (K-1)$. One of the variables in x_{iS} captures the geographical clustering of licenses in collection S. The term β_0 denotes the true parameter value. Under this parametrization, the marginal contribution of license l to collection S is $(x_{S\cup\{l\}} - x_S)\beta_0 + v_{il}$. Each bidder's stand-alone value for a license v_{il} is private information and hence is not observed by an econometrician.

Recall that H_{it} denotes the set of licenses bidder i is provisionally winning in the beginning of round t, and B_{it} the set of licenses on which bidder i places a bid during round t. Define A_{it} $= H_{it} \cup B_{it}$ as the union of these two sets. I call A_{it} bidder i's *portfolio* in round t. Bidder i could win A_{it} if the auction ends in round $t+1$. If the bidder is outbid on any of the licenses in H_{it}, then it may win a subset of A_{it}.[24]

If round t is late enough, bidder i may view A_{it} as one possible realization of H_{iT} given J_{it} and $(p_{lT}, p_{l'T}) = (p_{ilt},\ p_{il't})$. The difference between a particular H_{iT} and A_{it} gives rise to error. This error term can be viewed as expectational error that arises because bidder i does not know other bidders' bidding decisions and hence its own final collection. I let $\epsilon_{ill't}$ denote this non-structural error.[25] That is,

$$\Delta V_i(l, H_{iT}) - \Delta V_i(l', H_{iT}) = \Delta V_i(l, A_{it}) - \Delta V_i(l', A_{it}) + \epsilon_{ill't} \tag{1}$$

The difference between $\epsilon_{ill't}$ and zero should not affect the bidding decision since the bidder would have not viewed A_{it} as the "true" final collection. Therefore, it is reasonable to assume that $\epsilon_{ill't}$ is mean-independent of the variables known to the bidder when the bidding decision is made $E(\epsilon_{ill't}) = E(\epsilon_{ill't}|B_{it}, J_{it}) = 0$. The magnitude of this non-structural error term should shrink as the auction gets close to an end.

The presence of the unobserved bidder-license specific value v_{il} creates two problems. First, since v_{il} captures bidders' private information, it affects the bidder's decision, and therefore must have been selected from the subset of its possible values that would lead to the observed

[24] If multiple bidders place a bid on a license at the same bid amount, the FCC applies a random tie breaking rule to determine the standing high bidder for the license.

[25] I admit that this notation can be misleading. A more precise notation would be $\epsilon_{ill'H_{iT}}$.

decision. For example, bidder i will bid on l given the information available in round t only if $v_{il} \geq p_{ilt} - E((x_{A_{it}} - x_{A_{it}\setminus\{l\}})\beta_0 + \epsilon_{ilt} \big| J_{it})$ by BA1. Thus the expectation of the structural disturbance v_{il} conditional on the observed decision $l \in B_{it}$ can be different from its unconditional expectation, i.e. $E(v_{il}|l \in B_{it}, J_{it}) \neq E(v_{il})$.[26]

Second, the realizations of v_{il} can be correlated with the observed variables in the information set. For the case of spectrum licenses, Pop-MHz of a license is a major determinant of the stand-alone value of the license and also the license's marginal contribution to any bundle of licenses.

To account for these problems, I exploit the panel data structure by first-differencing out v_{il}. The bid data from a spectrum auction has a structure similar to panel data with the same bidder observed over many rounds, which corresponds to time periods in a panel data setting.

Let $\Delta x^S_{ill'} = x_S - x_{S\setminus\{l\}\cup\{l'\}} = (x_S - x_{S\setminus\{l\}}) - (x_{S\setminus\{l\}\cup\{l'\}} - x_{S\setminus\{l\}})$ denote the difference in the variables that capture the marginal synergy effects of license l and l' for collection S. Suppose bidder i was observed bidding on license l in a round, t, before it started bidding on license l'. If the bidder bid on license l' in round r after it stopped bidding on l, the difference of the marginal surpluses of the two licenses to the bidder's portfolio in round t and round r are, respectively,

$$\Delta x^{A_{it}}_{ill'}\beta_0 + v_{il} - v_{il'} + \epsilon_{ill't} \tag{2}$$

$$\Delta x^{A_{ir}}_{il'l}\beta_0 + v_{il'} - v_{il} + \epsilon_{ill'r} \tag{3}$$

Adding the two equations in 3 cancels out the unobserved stand-alone values of licenses l and l' and hence leads to

$$(\Delta x^{A_{it}}_{ill'} + \Delta x^{A_{ir}}_{il'l})\beta_0 + \epsilon_{ill't} + \epsilon_{ill'r}.$$

For notational simplicity, let $\Delta p_{ll't} \equiv (p_{lt} - p_{l't})$ denote the difference in the minimum acceptable bids of license l and license l'. BA2 leads to the following conditional moment inequality.

$$E\left[(\Delta x^{A_{it}}_{ill'} + \Delta x^{A_{ir}}_{il'l})\beta_0 + \epsilon_{ill't} + \epsilon_{ill'r} - (\Delta p_{ll't} - \Delta p_{ll'r}) \big| J_{it}\right] \geq 0 \tag{4}$$

The expectation is conditional on the information available in round t. Since $\epsilon_{ill't}$ is assumed to be mean-independent of the observables and does not affect the bidding decision, $E(\epsilon_{ill't}|J_{it}) = 0$. Also, the expectational error should also be mean independent to the past information set,

[26] To see the consequence of this selection problem, suppose an econometrician makes an assumption that the unconditional expected value of v_{il} is a function of observables. Due to selection on v_{il}, the (conditional) moment inequality implied by the behavioral assumptions may *not* be preserved at the true parameters if the inequality includes only the observables that approximate values.

i.e., $E(\epsilon_{ill'r}|J_{it}) = 0$.

The conditional moment inequality in (4) can be used to construct unconditional moment inequalities. Let a subscript j denote a particular bidder, license pair, round pair (i, l, l', t, r). Let $\mathbf{z}$ and $\mathbf{z}^K$ denote random variables whose realizations are $z_j \equiv (\Delta x_{ill'}^{A_{it}} + \Delta x_{il'l}^{A_{ir}})$ and $z_j^K \equiv (\Delta p_{ll't} - \Delta p_{ll'r})$, respectively. Let $\widetilde{\mathbf{z}}$ denote a random variable whose realization $\widetilde{z}_j = (\Delta x_{ill'}^{A_{it}}, \Delta p_{ll't})$ is contained in the information set J_{it} in round t. Set $\epsilon_j \equiv \epsilon_{ill't} + \epsilon_{ill'r}$.

$h^+(\cdot)$ and $h^-(\cdot)$ denote a real-valued non-negative and non-positive function, respectively. Since $E(h(\widetilde{\mathbf{z}}) \cdot \epsilon) = 0$ for any real-valued function $h(\cdot)$, the following unconditional moment inequalities hold.

Proposition 3

$$
\begin{aligned}
E\left[h^{sgn}(\widetilde{\mathbf{z}})\left(\mathbf{z}\beta_0 - \mathbf{z}^K\right)\right] &\geq 0 \;\; \textit{if } sgn = + \\
&\leq 0 \;\; \textit{if } sgn = -
\end{aligned}
\tag{5}
$$

Proof follows from PPHI. ■

$h^+(\cdot)$ and $h^-(\cdot)$ serve as instrumental variables that preserve the direction of the inequality sign implied by BA2. Note that I only consider observations in which I can cancel out the stand-alone values of licenses using the fixed effect approach. This implies that I use a subset of observations that can be derived from the behavioral assumptions to form the moment inequalities in (5).[27]

Let $Z_j = (z_j,\ z_j^K,\ \widetilde{z}_j)$ and $m_k^{sgn}(Z_j, \beta) = h_k^{sgn}(\widetilde{z}_j)(z_j\beta - z_j^K),\ k = 1,\ldots,K$ denote the kth moment function. A natural candidate for the non-negative function $h_k^+(\cdot)$ will be $h_k^+(\widetilde{z}_j^k) = 1(\widetilde{z}_j^k \geq 0)\widetilde{z}_j^k$ and for $h_k^-(\cdot)$, $h_k^-(\widetilde{z}_j^k) = 1(\widetilde{z}_j^k < 0)\widetilde{z}_j^k$ where $\widetilde{z}_j^k$ denotes the kth variable in $\widetilde{z}_j$. In this case, the number of moment inequalities equals $2K$. Define $s : \{+, -\} \to \{1, -1\}$ with $s(+) = 1$ and $s(-) = -1$. The identified set $\boldsymbol{\beta}_0$ is defined as the set of parameters satisfying all the moment inequalities. The estimate of $\boldsymbol{\beta}_0$ is the sample analogue of the identified set, i.e

$$
\boldsymbol{\beta}_N = \arg\min_{\beta \in \mathcal{B}} \sum_{k=1}^{Q} \left\| \min\left\{ s(sgn) \sum_{j=1}^{N} m_k^{sgn}(Z_j, \beta), 0 \right\} \right\|
$$

given the number of observations N. PPHI provides technical conditions under which the extreme point estimates are consistent. They also provide ways to approximate the limit

[27] There can be multiple pairs of licenses (l, l') that result from combining two observations from two rounds that cancel out the differences in the stand-alone values of the pair. For example, if bidder i was observed bidding on license l_1 in round t and then on license l_2 and l_3 in round r after it stopped bidding on l_1, pair of licenses (l_1, l_2) and (l_1, l_3) can be used for estimation. In this case, I randomly pick a pair so that each observation corresponds to a different (i, t, r).

distribution of the extreme point of each coefficient.

4.2 Second stage: Recovering the distribution of v_{il}

In the second stage, I attempt to estimate the distribution of stand-alone values of license v_{il} by parameterizing the distribution.

If bidder i bids on license l in round t, the net marginal contribution of the license to the bidder's portfolio at the price of the license in round t is

$$(x_{A_{it}} - x_{A_{it}\setminus\{l\}})\beta_0 + v_{il} + \epsilon_{ilt} - p_{lt} \tag{6}$$

Note that the behavioral assumptions do not imply the value of the equation in (6), which includes a realization of the expectational error, is either greater than or less than zero. However, the size of the expectational error term in (6) will get smaller as the auction nears an end. Recall that the expectational error arises from the possibility of being outbid on A_{it} eventually, so that the final collection is different from A_{it}. In later rounds, the chances of getting outbid on the licenses in A_{it} are small. Consequently, a bidder will believe the probability that it actually wins its portfolio is high.

I assume that if round t is late enough, I can ignore the non-structural error term ϵ_{ilt} and hence the following is true. If bidder i places a bid on license l in round $t \geq r_1$ where $r_1 \geq r_0$,

$$(x_{A_{it}} - x_{A_{it}\setminus\{l\}})\beta_0 + v_{il} - p_{lt} \geq 0 \tag{7}$$

If bidder i was observed being outbid on license l in round $\tau - 1$ and never bidding on a new license that may replace l as a special case of BA2,

$$(x_{A_{i\tau}} - x_{A_{i\tau}\setminus\{l\}})\beta_0 + v_{il} - p_{l\tau} \leq 0 \tag{8}$$

Hence, once the coefficients that govern the magnitude of the complementarities are estimated, the inequalities in (7) and (8) can be used to estimate the distribution of v_{il}. As (7) and (8) give a lower and upper bound of a realization v_{il} of a random variable $\mathbf{v}_l$, one can estimate bounding distributions of the distribution of $\mathbf{v}_l$ non-parametrically if there are enough observations that lead to inequalities similar to (7) and (8) for the same license l.[28] However, in a typical spectrum auction, only a few bidders bid on the same license and therefore it is not

[28] Another way of constructing an inequality that gives an upper bound for the stand-alone value of a license is to combine two inequalities from the modified BA2. Suppose bidder i had stopped bidding on l_1 before it started bidding on l. If the bidder bid on license l_1 in round r, the following inequality holds by the modified BA2:

$$\left(x_{A_{ir}} - x_{A_{ir}\setminus\{l_1\}\cup\{l\}}\right)\beta_0 + v_{il_1} - v_{il} \leq p_{l_1 r} - p_{lr} \tag{i}$$

feasible to estimate the distribution of the stand-alone value of each license non-parametrically. To resolve this problem, I make the following assumption.

Assumption 2 (Distribution of the stand-alone value) *v_{il} is drawn independently over licenses and bidders from an normal distribution $\Phi(\mu_{il}, \sigma_{il}^2)$ with $\mu_{il} = \gamma_0^1 z_{il}^1$ and $\sigma_{il} = \gamma_0^2 z_{il}^2$.*

The distribution of the stand-alone values of a license is normal with a unknown mean and variance. The mean and standard deviation are assumed to be functions of bidder-license specific characteristics. There is another reason why this assumption is necessary, beyond a limited number of observations for each license: As I only consider "small" licenses, I have to extrapolate the distribution of the stand-alone value of a "large" license from that of a "small" license. Note that this assumption was not needed for the first stage estimation. As I cancel out the stand-alone values, the first stage is free from any parametric assumption on the distribution of v_l.

As only a set to which the true parameter β_0 belongs is identified by the model, for each $\beta \in \boldsymbol{\beta}_0$, there will be a corresponding (7) and (8). Define $\underline{\beta}_{S,l}^0 = \arg\min_{\beta \in \boldsymbol{\beta}_0} \left(x_{S\cup\{l\}} - x_{S\setminus\{l\}}\right)\beta$ and $\overline{\beta}_{S,l}^0 = \arg\max_{\beta \in \boldsymbol{\beta}_0} \left(x_{S\cup\{l\}} - x_{S\setminus\{l\}}\right)\beta$. Then, I obtain the following inequalities.

$$v_{il} \leq p_{lt} - (x_{A_{it}} - x_{A_{it}\setminus\{l\}})\underline{\beta}_{A_{it},l}^0 \tag{9}$$

$$v_{il} \geq p_{l\tau} - (x_{A_{i\tau}} - x_{A_{i\tau}\setminus\{l\}})\overline{\beta}_{A_{i\tau},l}^0 \tag{10}$$

Let a subscript q denote a distinct bidder-license pair (i, l). Letting u_q denote the right hand side of (9) and d_q, of (10) gives

$$d_q \leq v_q \leq u_q$$

Suppose the true parameters $\gamma_0 = (\gamma_0^1, \gamma_0^2)$ are known. I can standardize u_q, v_q and d_q by subtracting the mean μ_q of the distribution of $\mathbf{v}_q$ from each of them and then dividing it by the standard deviation σ_q of $\mathbf{v}_q$:

$$\frac{d_q - \mu_q}{\sigma_q} \leq \frac{v_q - \mu_q}{\sigma_q} \leq \frac{u_q - \mu_q}{\sigma_q} \tag{11}$$

Let $\overline{\mathbf{u}}$ denote the standardized random variable whose qth realization is $\overline{u}_q = \dfrac{u_q - \mu_q}{\sigma_q}$. I define

Adding (i) and (8) gives

$$\beta_0 \left(x_{A_{ir}} - x_{A_{ir}\setminus\{l_1\}\cup\{l\}} + x_{A_{i\tau}} - x_{A_{i\tau}\setminus\{l\}}\right) + v_{il_1} \leq p_{l_1 r} - p_{lr} + p_{l\tau}$$

A lower bound for license l_1 can be constructed in a similar way.

$\overline{\mathbf{d}}$ and $\overline{\mathbf{v}}$ similarly.

(11) holds for any μ_q and σ_q. Since the random variable $\mathbf{v}_q$ follows a normal distribution $\mathbf{v}_q \sim \Phi(\mu_q, \sigma_q^2)$, the distribution of $\overline{\mathbf{v}}$ will follow the standard normal distribution at the true parameter $\gamma_0 = (\gamma_0^1,\ \gamma_0^2)$. Therefore, at the true parameter $\gamma_0 = (\gamma_0^1,\ \gamma_0^2)$, the distribution of $\overline{\mathbf{u}}$ stochastically dominates the standard normal distribution and the distribution of $\overline{\mathbf{d}}$ is stochastically dominated by the standard normal distribution.

Let $\overline{G}$ denote the distribution of $\overline{\mathbf{d}}$ and $\underline{G}$ the distribution of $\overline{\mathbf{u}}$. Given $\boldsymbol{\beta}_0$, at the true parameter γ_0, $\underline{G}(t, \gamma_0, \boldsymbol{\beta}_0) \leq \widetilde{\Psi}_{t_0,t_1}(t) \leq \overline{G}(t, \gamma_0, \boldsymbol{\beta}_0)$ for all $t \in [t_0, t_1]$ where $\widetilde{\Psi}_{t_0,t_1}$ denotes the standard normal distribution truncated above t_0 and below t_1. Note that the values of v_q that satisfy the inequalities in (9) and (10) must have been selected from the whole support of its distribution. For example, if a bidder's stand-alone value for a license is very large, the bidder will not stop bidding for the license, and hence its upper bound will never be observed. The fact that each random variable follows a truncated normal distribution will cause the standard normal distribution constructed from the variables to be truncated. Therefore, the identified set Γ_0 is defined as:

$$\boldsymbol{\Gamma}_0 = \{\gamma \in \boldsymbol{\Gamma} | \underline{G}(t, \gamma_0, \boldsymbol{\beta}_0) \leq \widetilde{\Psi}(t) \leq \overline{G}(t, \gamma_0, \boldsymbol{\beta}_0)\ \forall\ t \in [t_0, t_1]\ \}$$

Following Manski and Tamer (2002), $\gamma \in \boldsymbol{\Gamma}$ minimizes the following criterion function if and only if $\gamma \in \boldsymbol{\Gamma}_0$:

$$Q(r, \eta) = \int_{t_0}^{t_1} \left(\min \left[\overline{\phi}(t; \gamma, \eta), 0 \right]^2 + \min \left[\underline{\phi}(t; \gamma, \eta), 0 \right]^2 \right) d\widetilde{\Psi}(t) \tag{12}$$

where $\overline{\phi}(t; \gamma, \boldsymbol{\eta}) \equiv \overline{G}(t; \gamma, \boldsymbol{\beta}) - \widetilde{\Psi}_{t_0,t_1}(t)$ and $\underline{\phi}(t; \gamma, \boldsymbol{\eta}) \equiv \widetilde{\Psi}_{t_0,t_1}(t) - \underline{G}(t; \gamma, \boldsymbol{\beta})$ where $\boldsymbol{\eta} = (\boldsymbol{\beta}, t_0, t_1)$.

To estimate Γ_0, define $\widehat{\underline{\beta}}_{S,l} = \underset{\beta \in \boldsymbol{\beta}_N}{\arg\min} \left(x_{S \cup \{l\}} - x_{S \setminus \{l\}} \right) \beta$ and $\widehat{\overline{\beta}}_{S,l}$ similarly. An estimate $\widehat{\overline{d}}_q$ of $\overline{d}_q$ can be obtained by replacing $\underline{\beta}^0_{A_{it},l}$ with $\widehat{\underline{\beta}}_{A_{it},l}$ and an estimate $\widehat{\overline{u}}_q$ of $\overline{u}_q$ can also be similarly obtained. Let $\widehat{\overline{G}}(\cdot; \boldsymbol{\beta}_N, \gamma)$ denote the empirical distribution function of $\widehat{\overline{d}}_q$ and $\widehat{\underline{G}}(\cdot; \boldsymbol{\beta}_N, \gamma)$ of $\widehat{\overline{u}}_q$.

Let n_2 denote the number of observations used for the construction of the empirical distributions $\widehat{\underline{G}}(\cdot; \boldsymbol{\beta}_N, \gamma)$ and $\widehat{\overline{G}}(\cdot; \boldsymbol{\beta}_N, \gamma)$. Define $n = (N, n_2)$. The second stage estimator is obtained by minimizing the sample analogue of (12):

$$Q_{n_I}(\gamma; \widehat{\eta}_n) = \frac{1}{n_I} \sum_{k=1}^{n_I} \left(\min \left\{ \widehat{\overline{\phi}}_{n_2}(t_k; \gamma, \widehat{\eta}_n), 0 \right\}^2 + \min \left\{ \widehat{\underline{\phi}}_{n_2}(t_k; \gamma, \widehat{\eta}_n), 0 \right\}^2 \right) \tag{13}$$

To establish consistency of the estimator, I assume:

Assumption 3 *(i) The* $t_1, t_2, \ldots, t_{n_I} \in [t_0, t_1]$ *are i.i.d. draws from* $\widetilde{\Psi}$.

(ii) As $N \longrightarrow \infty$*, both* $n_2, n_I \longrightarrow \infty$

(iii) The set $\mathbf{\Gamma} \subset R^M$ *is compact*

(iv) $\widehat{t}_0$ *and* $\widehat{t}_1$ *are a consistent estimate of* t_0 *and and* t_1*, respectively.*

Define

$$\Gamma_N = \left\{ \gamma \in \Gamma \middle| \widehat{Q}_{n_I}(\gamma; \widehat{\eta}_n) \leq \min_{r \in \Gamma} Q_{n_s}(\gamma; \widehat{\eta}_n) + \epsilon_N \right\} \tag{14}$$

and let $\rho(\Gamma, \Gamma') \equiv \sup_{\theta \in \Gamma} \inf_{\theta' \in \Gamma'} |\theta - \theta'|$ measure the distance between two sets $\Gamma, \Gamma' \subset \mathbb{R}^M$. The following proposition establishes consistency of the estimator.

Proposition 4 *Under Assumption 3,* $\rho(\Gamma_N, \Gamma_0) \xrightarrow{P} 0$.

Furthermore, if $\sup_{\gamma \in \mathbf{\Gamma}} \left| \widehat{Q}_n(\gamma; \widehat{\eta}_n) - Q(\gamma; \eta) \right| \Big/ \epsilon_N$ *,then* $\rho(\Gamma_0, \Gamma_N) \xrightarrow{P} 0$

Proof See Appendix B. ■

5 Application to AWS-1

This section describes the AWS-1 and provides some summary statistics. I also discuss the parametrization choices that I make when applying the estimation technique to AWS-1.

5.1 AWS-1

The AWS-1 auction began in June 2006 and ended in September 2006 after 161 rounds. The auction was designed to provide an additional spectrum ranging from 1710-1755 MHz and 2110-2155 MHz to wireless service providers wishing to offer a variety of wireless services, including Third Generation ("3G") mobile broadband and other advanced wireless services. The FCC offered 1,122 licenses, with 13.7 billion Pop-MHz, for sale and 168 bidders participated in the auction. The auction generated \$13.7 billion of revenue for the US Treasury after 104 bidders won 1,087 licenses. The average price per Pop-MHz was \$0.533.

Table 1 presents the licenses offered for sale in detail. The FCC divided the 90 MHz of spectrum into 6 frequency blocks, listed as A through F. To define coverage of a license in each frequency block, the FCC used various definitions of geographical service areas. For Block A, the FCC used the market definition of Cellular Market Area (CMA)s, which consist of 306 MSA (Metropolitan Statistical Area)s and 428 RSA (Rural Service Area)s on the basis of the 1980 Census. For block B and C, the FCC applied a service area map called Basic Economic

Table 1: Licenses offered for sale in AWS-1

No. of licenses offered (sold)	Total Pop-MHz[a]	Revenue[b]	Price per Pop-MHz
1,122 (1,087)	25.706	13.7	\$0.533

Geographic Service Areas	Blocks (size of bandwidth)	No. of licenses offered for each block
MSA or RSA	A(20)	734
Basic Economic Area (BEA)	B(20), C(10)	176
Regional Economic Area Grouping (REAG)	D(10), E(10), F(20)	12

[a]: in billions. [b]: in billion dollars

Area (BEA), which divides the US and its territories into 176 areas. For the remaining blocks, Regional Economic Area Grouping (REAG) which includes 8 service areas in the continental US and 4 in its territories was used to define coverage of spectrum licenses.

Out of 168 bidders who made an upfront payment for eligibility, 164 bidders actually participated in the auction placing at least one bid throughout the auction. Small bidders were awarded discounts (bidding credits) on their winning bids: 15% for small business and 25% for very small business.[29] Out of 164 bidders, 44 were awarded 15% bidding credits and 54 bidders 25% bidding credits.

Throughout 160 rounds, 16,197 total bids were placed on 1,092 licenses, averaging 14.8 bids on a license. There were 159 licenses that received only one bid. Bid withdrawals were not frequent—only 11 bidders made 26 bid withdrawals in total, resulting in five licenses unsold and eight licenses sold but subject to the withdrawal payments.

The AWS-1 auction opened with relatively high minimum opening bids compared to previous auctions. The minimum opening bids were as large as 45% of the final prices on average with a median of 32.6%. However, as many licenses' winning bids were much larger than their minimum opening bids, the sum of the opening bids was only 8.4% of the sum of the final winning bids. Although the FCC provides nine acceptable bids for each license each round, most of the bids were exactly the minimum acceptable bids accounting for 98.5% of the total bids placed. Only 1.1% (240 bids) exceeded the minimum acceptable bids by more than 5%, and 31 bids out of these 240 jump bids became the winning bids.

Table 2 presents the 4 biggest winners in AWS-1 with an emphasis on the asymmetry amongst winners. The 4 biggest winners accounted for 71% of the total units of Pop-MHz

[29]Very small business and small business were defined as business whose average gross revenue for the preceding three years did not exceed \$15 million and fell in between \$15 million and \$40 million, respectively.

Table 2: Major winning bidders in AWS-1

	T-Mobile	SpectrumCo	Verizon	Cingular	Top 4 total	Others
Pop-MHz[a]	6.64	5.27	3.84	2.44	18.18	7.36
Net payment[a]	4.18	2.38	2.81	1.33	10.70	3.00
Price per Pop-Mhz	0.63	0.45	0.73	0.55	0.59	0.41
Percentage of Pop-MHz	25.99%	20.62%	15.04%	9.54%	71.19%	28.81%

[a]: in billions.

sold and 78% of the total revenue. T-Mobile was the biggest winner in terms of both units of Pop-MHz contained in the winning collection accounting for 25.99% of the total amount of Pop-MHz sold and the net payment. Only 11 bidders, including the 4 biggest winners, won more than 1% of the total Pop-MHz sold and the rest, 93 bidders won 0.64% of Pop-MHz sold on average.

The activity requirement for AWS-1 was 80% of the bidder's eligibility until round 30, from which point it then became 95%. As a result of the activity rule, the number of qualified bidders who were holding a positive amount of eligibility gradually decreased. At about round 30, only 136 bidders out of the originally qualified 168 bidders remained qualified.

5.2 Parametrization: complementarities

As the geographic licensing schemes used to divide the US and its territories are different for each frequency block, the geographic size of a license exhibits a large amount of variation across frequency blocks. Given this, it is not clear what the stand-alone value means for a license that covers a large area that can be covered by a collection of several small licenses. For example, if there are positive complementarities among CMA licenses that belong to the same BEA, the stand-alone value of the BEA license must include these complementarities.

To be consistent across frequency blocks, I let "a license" denote a CMA license. A large license, such as a BEA or a REAG license, is defined as the "collection" of the CMA licenses that belong to the large area.[30] Formally, the stand-alone value is defined as follows:

$$v_{il} = V_i(\{l\}) \text{ if } l \text{ is a CMA license}$$

Therefore, the value of a BEA or a REAG license is

[30] To be precise, the building blocks for BEAs and REAGs are not CMAs but counties. Counties are also the basic building blocks for CMAs. It occurred therefore, though not often, that two counties that belong to one common CMA belong to two different BEAs or even two different REAGs. Whenever this occurs, I adjust the large license to contains the right portion of the CMA license.

$$V_i(\{l\}) = V_i(C_l) = \sum_{l' \in C_l} v_{il'} + k_{iC_l} \text{ if } l \text{ is either a BEA or a REAG license}$$

where $C_l = \{\, l' \in \mathcal{T} |\ l$ covers the CMA area associated with $l'\}$ and k_{iC_l} denotes the complementarities amongst the CMA licenses that belong to C_l.

With this definition, I convert any collection of licenses into a collection of CMA licenses with the bandwidth size and population adjusted appropriately. For example, if a bidder's collection contains one 10 MHz BEA license which covers CMA160 and some part of CMA360, and a 20 MHz CMA license, say CMA361, the converted collection will be the collection of 20 MHz CMA361, 10 MHz CMA160, and 10 MHz CMA360 that contains only the right portion of the whole population of CMA360. A collection of licenses hereafter refers to the converted collection.

For complementarities, I construct two variables. The first variable captures synergies achieved through common ownership of geographically close licenses. Geographic complementarities may arise due to the presence of the minimum investment requirement for infrastructure deployment, and hence economies of scale, or advantages in advertisement and the licensee's market position as a network operator with extensive and continuous geographic coverage. The presence of geographic complementarities is well documented by several empirical studies. The pairwise geographic synergies between two licenses will be positively related to the capacity-adjusted population size (Pop-MHz) of the two areas and negatively to the distance between the two areas as noted by Moreton and Spiller (1998). The variable is constructed as follows:

$$\text{geocomple}_S = \sum_{l \in S} \sum_{l' \in S \backslash \{l\}} \frac{\text{Pop-MHz}_l (\text{Pop-MHz}_{l'})^{1/2}}{d_{l,l'}^2}$$

where $S_l = \{\, l' \epsilon S \backslash \{l\} |\, l$ and l' belong to the same state$\}$. $d_{l,l'}$ denotes the distance between the area associated with license l and the area associated with license l'. The distance between two CMA areas, say l and l', was measured as the minimum distance between a county in l and a county in l'. This variable is more conservative than the one used in Bajari and Fox (2007). The pairwise complementarities between two licenses l and l', $\frac{\text{Pop-MHz}_l (\text{Pop-MHz}_{l'})^{1/2}}{d_{l,l'}^2}$, is positive only if the two licenses belong to the same state.

This conservatism is appropriate due to the following reasons. First, many large bidders in AWS-1 were already operating a wireless network regionally or nationally before the auction. For these bidders, the presence of the minimum investment requirement for infrastructure deployment, and hence economies of scale is not likely to be a factor that constitutes complementarities, and global synergy effects should be addressed in combination with their current holdings of other spectrum licenses. Second, small bidders such as rural telephone companies tend to bid on several licenses contained in one or two states. This indicates that a firm may not need more than several CMA licenses to achieve an efficient scale of operation. Based on these considerations, I conclude that the more important factor for geographic synergies is advantages in advertising, market position and network management. These are likely to be achieved at the state-level where wireless industry regulation is shared.

The second collection-specific variable is constructed by summing the pairwise complementarities that measure how close two CMAs are in terms of travelers over licenses in the collection.

$$\text{travelcomple}_S = \sum_{l \in S} \sum_{l' \in S \setminus \{l\}} t_{l,l'} \text{Pop-MHz}_l (\text{Pop-MHz}_{l'})^{1/2}$$

where $t_{l,l'} = \dfrac{\text{Number of passengers with origin } l \text{ and destination } l'}{\sum\limits_{l' \in T} \text{Number of passengers with origin } l \text{ and destination } l'}$ measures the relative importance of an area l' as destination of trips from area l.

This was constructed using T-100 Domestic Market Airline Traffic Data for the calendar year 2005. The data includes passenger counts enplaned at the origin airport and deplaned at the destination airport reported by U.S. air carriers operating between airports located within the boundaries of the United States and its territories. Recognizing there are multiple airports within a CMA, I aggregate passenger counts at the CMA level. This variable captures synergies that may arise from serving traveling customers without charging extra for roaming service. It also fills in the global synergy effects left out by the first variable that only accounts for state-level local synergy effects.

Note that the complementarities are common across bidders by the construction of the complementarity variables. The reason for this specification is because I do not have data that captures bidder heterogeneity relevant to synergy effects. The estimation procedure allows complementarity variables to differ across bidders. This specification, however, is plausible since main sources of the complementarities are likely to be properties of wireless service production technology, and consumers' strong preferences for seamless services as discussed above.[31] This assumption implies the sources of variation in the collection of licenses each bidder pursues are random draws for stand-alone values of licenses that are privately observed.

[31] Brusco and Lopomo (2002) argue that what matters for signaling incentives is not the existence of complementarities but whether the complementarities are common across bidders or not. Intuitively, as common complementarities can be fully competed away in a competitive equilibrium, the existence of common complementarities does not destroy incentives to collude. By assuming commonness in the magnitude of complementarities, we lose an opportunity to test, for example, whether variability in the complementarities is large enough to prevent a rise of a collusive equilibrium.

Table 3: Explanatory power of a complementarity variable on the last bids

Round at which the first bid is placed	No. of obs.	$R^2_{(1)}$	$R^2_{(2)}$	$R^2_{(2)} - R^2_{(1)}$	$(R^2_{(2)} - R^2_{(1)})/R^2_{(1)}$
all (>0)	2275	0.586	0.598^a	0.012	0.020
>10	1702	0.634	0.655	0.021	0.033
>20	1206	0.555	0.648^b	0.093	0.168
>30	670	0.718	0.818	0.100	0.139
>40	362	0.667	0.822	0.155	0.232
>50	248	0.617	0.792	0.175	0.284

All licenses except the CMA licenses over the US territories were used for regression. All the coefficients are statistically significant and positive except the first and third regression of (2) where ^{a}: the coefficient for travelcomple is not significant and ^{b}: the coefficient for geocomple is negative.

5.2.1 Bidding above the stand-alone values

Recall that BA2, per se, does not require a bidder to bid above the stand-alone value for a license before it stops bidding on it. Even in the case where the license's marginal contribution to the bidder's portfolio is greater than the license's stand-alone value, the bidder does not need to bid up to the marginal contribution before it drops the license. However, either if the bidder never bids above the stand-alone value of a license, or if it never considers the marginal synergy effect of the license on its portfolio, BA2 will not be informative as to the magnitude of the complementarities. Table 3 provides evidence that suggests the maximum bid submitted by a bidder on a particular license is positively correlated to the license's marginal synergy effect on the bidder's portfolio. Furthermore, this relationship is more pronounced in later rounds.

Table 3 compares two sets of regressions in which each regression considers only a subset of observations. Let b_{il}^r denote the maximum bid amount placed by bidder i on license l. Let r denote the round when the maximum bid was placed. Note that there is only one observation for a bidder-license pair because the round when the bidder bid on the license at the maximum bid amount is unique for each pair.

The first set (1) regresses the maximum bid submitted by a bidder for a license on a constant and the license's Pop-MHz:

$$(1)\ b_{il}^r = \alpha_{(1)} + \beta_{(1)}\text{Pop-MHz}_l + \varepsilon_{il}$$

The second set (2) regresses the maximum bid on a constant, the license's Pop-MHz, and two variables that capture the marginal synergy effect of the license on the bidder's portfolio in the round when the maximum bid was placed:

$$(2)\ b_{il}^r = \alpha_{(2)} + \beta_{(2)}\text{Pop-MHz}_l + \gamma^{geo}(\text{geocomple}_{A_{ir}} - \text{geocomple}_{A_{ir}/\{l\}}) \\ + \gamma^{tra}(\text{travelcomple}_{A_{ir}} - \text{travelcomple}_{A_{ir}/\{l\}}) + \widetilde{\varepsilon}_{il}$$

To see whether the license's marginal synergy effect on the bidder's portfolio has a larger effect on the bidder's bidding decision in later rounds, I divide the observations into several groups by the round of the bidder's first bid. Except the first row, each regression includes a subset of the observations depending the first-bid rounds.

The regression results in Table 3 suggest that the marginal synergy effect of a license to a bidder's portfolio has a positive effect on the maximum bid submitted by the bidder on the license. They also show that the additional explanatory power of the marginal synergy effect variables on the maximum bid amounts, $R^2_{(2)} - R^2_{(1)}$, tends to grow larger as the regression includes only the licenses chosen by bidders in later rounds.

This result is consistent with my argument that in late rounds, bidders consider their portfolios as the set of licenses it is likely to win. Otherwise, the marginal synergy effect of the license on the bidder's portfolio would not have any information on the maximum bid amounts. Also, it supports the fundamental assumption of this paper that bidders' decisions to bid on individual licenses reveal information on the magnitude of the complementarities.

5.3 Parameterization: distribution of stand-alone values

For the second stage estimation, I divide the bidders into two groups: local bidders and global bidders. As the FCC offered licenses in various sizes in AWS-1, participating firms also varied in size and business plan. It is hard to imagine that national wireless carriers such as Verizon and T-Mobile or the joint venture SpectrumCo, which includes Comcast, Time Warner and Cox, and Sprint Nextel, would have business plans similar to a local telephone company or a local Internet service provider with an average gross revenue around $20 million. To account for this obvious asymmetry, I call a bidder a local bidder if either the bidder classified itself as a rural telephone company in the application form submitted to the FCC or if the bidder has bid for a smaller collection of licenses that cover a confined region of up to five states. Out of 164 bidders who placed a bid, 72 identified themselves as a rural telephone company and 44 bid on only one or two licenses confined in one state throughout the auction. These bidders are classified as local bidders. Besides these bidders, 19 bidders who bid on licenses within several nearby states were also counted as a local bidder.[32] Table 4 presents some summary statistics of the two bidder groups. It shows that the local bidder bidders, on average, began with a lower level of eligibility, compared to the global bidders. Many of them were also awarded some bidding credits.

Following Section 4, the mean of the distribution of a license's stand-alone value is assumed

[32] Four bidders who bid on only one to three licenses confined to a small area in the US territories, such as Puerto Rico, are classified as neither a local nor a global bidder, and hence are not considered.

Table 4: Global vs. local bidders

Type	No. of bidders	Bidding credit	Initial eligibility ratio	Activity (billion Pop-MHz)
Global	25	0.082	0.123	1.29
Local	135	0.128	0.00048	0.009

Initial eligibility ratio of bidder i is defined as the ratio of bidder i's initial eligibility to the amount of bidding units required to purchase all the licenses offered for sale in AWS-1.

to be a linear function of variables that affect the value. The mean and standard deviation of the distribution of stand-alone values for license l to bidder i is parameterized as:

$$\begin{aligned} \mu_{il} &= \gamma_1^j \text{Pop-MHz}_{il}^{\text{e}} + \gamma_2^j \text{Dist}_{il} + \gamma_3^j \text{Area}_l \\ \sigma_{il} &= \gamma_\sigma^j \text{Pop-MHz}_{il}^{\text{e}} + \vartheta^j, \;\; j = \text{local, global} \end{aligned} \tag{15}$$

A positive coefficient for Pop-MHz captures the marginal increase in the average value of a license due to the increase in the number of potential subscribers in the market, multiplied by the bandwidth of the license. The bandwidth of a license corresponds to the capacity of the license which determines the number of simultaneous phone calls it can handle. The Pop-MHz of a license is expected to be an important determinant of the value of the license.

$\text{Pop-MHz}_{il}^{\text{e}}$ denotes eligibility-adjusted Pop-MHz. Admittedly, the covariates in (15) are not enough to account for various business plans of the AWS-1 participants. This will result in a large variance estimate in (15). This is especially the case for the local bidder group. Although they are grouped together as local bidders, the group includes very small bidders who bid only on one license throughout the auction as well as others who have won licenses over several states. To account for this difference, $\text{Pop-MHz}_{il}^{\text{e}}$ instead of Pop-MHz is used for estimation. $\text{Pop-MHz}_{il}^{\text{e}}$ is defined as $\text{Pop-MHz}_{il}^{\text{e}} \equiv \text{Pop-MHz}_l \cdot elig_i$ where $elig_i = \log(1 + 10^4 \cdot \text{eligibility ratio}_i)/\,4$ if i is local and 1 otherwise. The numerator of 4 ensures that a local bidder's $\text{Pop-MHz}_{il}^{\text{e}}$ is less than Pop-MHz_l.[33]

Dist_{il} denotes the distance in meters between license l and the CMA area of bidder i's address on the application form submitted to the FCC. If a local bidder values a license close to its business address more than a distant license, the coefficient should be negative. I assume the coefficient for this variable is zero for the global bidder group. This assumption is plausible because the global bidders, which include nationwide cellular phone carriers, will not particularly value the licenses close to their headquarters. If the coefficient of Dist_{il} is negative, the model will predict that a local bidder will pursue a smaller collection compared to a global

[33] The specific functional form of an eligibility ratio, $\log(1+10^4 \cdot \text{eligibility ratio}_i)$, is chosen to reduce variation of $\text{Pop-MHz}_{il}^{\text{e}}$ resulting from large differences in local bidders' eligibility ratios.

bidder on average even if all other coefficients are the same.

I attempt to capture the cost side effect on a license's value by including Area_l. The number of base transceiver stations for a region depends on the density of subscribers, the region's area and terrain. A base station is limited in the number of simultaneous phone calls it can deal with and the area it can cover. Therefore, the area of a region will affect the wireless service production cost for the market.[34]

The variables in (15) are expected to be related to a license's value positively, or negatively, depending on what each variable is intended to capture. However, if these variables are also correlated to competition level, then it is not clear what sign to expect since we are missing the data on competition level for each market. I allow the parameters in the mean and standard deviation to differ depending on whether the CMA license is associated with MSA or RSA for the global bidder group. I could not do so for the local bidders because of two few second stage observations.

5.4 Choice of moments

To estimate the coefficient for the complementarity variables in the first stage, I use six moment conditions, each of which holds as an inequality. Let S denote a set of licenses. Suppose that bidder i bids on license l in round t, but not on license l'.

Define $\Delta\text{geocomple}_{ll'}^{S} \equiv \text{geocomple}_{S\cup\{l\}} - \text{geocomple}_{S\setminus\{l\}\cup\{l'\}}$, and $\Delta\text{travelcomple}_{ll'}^{S}$ similarly. Following the notation in Section 4.1, let $\widetilde{z}_1$, $\widetilde{z}_2$, and $\widetilde{z}_3$ denote the vectors of $\Delta\text{geocomple}_{ill'}^{A_{it}}$, $\Delta\text{travelcomple}_{ill'}^{A_{it}}$ and $\Delta p_{ll't}$, respectively.

The set of the instrumental variables, given as $h^{+}(\cdot)$ and $h^{-}(\cdot)$ in (5), that I used to construct the six moment conditions are $\{(\widetilde{z}_j \geq 0,\ \widetilde{z}_3 \geq 0)\widetilde{z}_j,\ (\widetilde{z}_j \geq 0,\ \widetilde{z}_3 < 0)\widetilde{z}_j,\ (\widetilde{z}_j < 0)\widetilde{z}_j\}_{j=1,2}$. These instrumental variables are chosen to ensure that the estimated set satisfying all the moment inequalities is non-empty. The estimated set resulting from more natural moment conditions constructed by $\{(\widetilde{z}_j \geq 0)\widetilde{z}_j,\ (\widetilde{z}_j < 0)\widetilde{z}_j\}_{j=1,2,3}$ is also non-empty, and is smaller. However, I use the formal moment conditions. This choice ensures that $\widehat{\overline{G}}$ stochastically dominates $\underline{\widehat{G}}$ because the formal moment conditions yield a slightly bigger set, which contains the estimated set from the latter moment conditions.[35] Given the finite number of observations, I conclude that this conservative approach to estimate the set containing the true parameters is appropriate.

[34]The current specification does not include the population density, the interaction term of population and area as an determinant of the value of a license. The wireless service technology predicts that while the population density is more important in determining the number of base stations for most urban regions, the area is more important in rural regions.

[35]A violation of the stochastic relationship between $\widehat{\overline{G}}$ and $\underline{\widehat{G}}$ does not imply misspecification, given the finite number of observations.

Table 5: First Stage Estimation Results

	Interval Estimate	Min.	Max.
geocomple	[4.92, 67.25]	4.92	67.25
simulated 95% *CI*	[2.63, 127.53]	[2.63, 26.27]	[6.65, 127.53]
conservative 95% *CI*	[−10.48, 129.69]	[−10.48, 80.78]	[−97.52, 129.69]
travelcomple	[0.008, 0.123]	0.008	0.123
simulated 95% *CI*	[0.005, 0.131]	[0.005, 0.010]	[0.0144, 0.125][a]
conservative 95% *CI*	[−0.053, 0.237]	[−0.053, 0.011]	[−0.187, 0.236][a]

All figures are in thousandths (10^{-3}). The simulated CIs are based on 10,000 simulation draws. [a] reports the 90% confidence interval. To simulate the distributions that dominates and is dominated by the true distribution, I used the same set of moment inequalities.

I choose round 31 as r_0 from which the inequalities in BA1 and BA2 are satisfied. I choose round 31 because the activity requirement had increased to 95% from round 31, and therefore bidders could not delay bidding for licenses that they desire. The regression results in Table 3 also support this choice.

I make several further restrictions on the pair of licenses (l, l') for which the revealed preference inequalities in BA2 should hold. First, I restrict the pair of licenses (l, l') to be the same type of CMA licenses. That is, if license l is a MSA license, l' should be also a MSA license, and if l is a RSA license, l' should be also a RSA license. This restriction reduces the number of observations, but reinforces the substitution motivation that BA2 tries to capture.

Second, choosing between license l and license l', the decision should not be influenced by considerations of current and future eligibility. I require that bidder i's eligibility level in round t does not prevent bidder i from bidding on license l'. I further require that if bidder i bids on license l', this alternative decision should not violate the activity requirement. This restriction controls for the fact that the behavioral assumptions do not explicitly account for the activity rule.[36]

6 Estimation Results

Table 5 presents the first stage estimation results. The coefficients in the complementarities are both positive. The first panel provides the interval estimate and inference for the true parameter β_0 for each coefficient while the second panel provides for the the extreme values of each coefficient. Since zero lies outside of the 95% confidence interval (CI) for each coefficient, the

[36] Suppose that bidder i bids on license l, but not on license l' because either the bidder was not eligible for license l' or it could have not met the activity requirement by bidding on license l'. In this case, the premise of BA2 is not satisfied because the alternative decision would change the distribution of the set of licenses that maximizes bidder i's surplus at each final price vector.

Table 6: Second Stage Estimation Results: Local Bidders

Mean (μ)	
Pop-MHz^e	[0.300, 0.434]
Land Area (mi^2)	[-6.031, 3.398]
Distance From Business Address (m)	[-1.036, -0.657]
Standard Deviation (σ)	
Pop-MHz^e	[0.124, 0.232]
ϑ	[2.68×10^5, 3.16×10^5]

The estimated set was found by solving the minimization problem for randomly drawn initial guesses using a regular search and simulated annealing algorithm for about 2000 times. 79 solutions were found. After this work was done, I ran a constrained minimization algorithm repeatedly until the minimum and maximum value of each parameter does not change. In total 255 solutions were found. A confidence interval based on a subsampling method is in progress.

null hypothesis that any of these complementarity variables has no effect on bidder valuations is rejected at the 5% significance level when tested using the *simulated* distribution. I also reject the hypothesis that the minimum and the maximum of each coefficient is zero at the 5% significance level with one exception. The maximum of the travel complementarities coefficient is only significant at the 10% level.

I fail to reject these hypotheses when testing using the more conservative 95% confidence interval. As for the inference of the true parameters, considering the fact that the simulated confidence interval for each parameter is also conservative, the statistical significance test for each parameter based on the simulated confidence interval is not likely to suffer from over-rejection.[37]

Table 6 presents the second stage estimation result for local bidders. It shows that a one-unit increase in the Pop-MHz of a license increases the mean of a local bidder's stand-alone value for the license by 30 to 43 cents. The increase is smaller for bidders with smaller eligibility ratios $elig_i$. This result says that bidders with small initial eligibility do not value a large Pop-MHz license as much as bidders with large initial eligibility. However, this result should not be read as a causality relationship. It should be understood that the eligibility term in Pop-MHz^e accounts for some bidder heterogeneity that the other variables fail to capture.

The result also shows that the further away the license's market is from a local bidder's

[37] There are three sources of conservatism in the construction of the conservative CI. The two sources of conservatism that the simulated CI and the conservative CI share are due to the two inequalities in $\Pr(\beta_0 \in [\widehat{q}_{\alpha/2}, \widehat{q}_{1-\alpha/2}]) \geq \Pr([\underline{\beta}_0, \overline{\beta}_0] \in [\widehat{q}_{\alpha/2}, \widehat{q}_{1-\alpha/2}]) \geq 1 - \Pr(\underline{\beta}_0 < \widehat{q}_{\alpha/2}) - \Pr(\overline{\beta}_0 > \widehat{q}_{1-\alpha/2})$. The third source of conservatism that only the conservative CI comes from the fact that its construction makes sure that $\Pr(\underline{\beta}_0 < \widehat{q}_{\alpha/2}) \leq \alpha/2$ and $\Pr(\overline{\beta}_0 > \widehat{q}_{1-\alpha/2}) \leq \alpha/2$. See PPHI *pp 36-37* for more discussion.

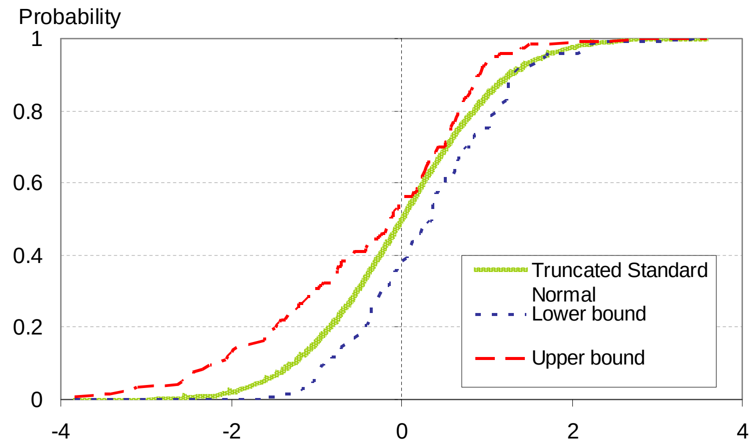


Figure 2: Second Stage Estimation: Local Bidders

Note: The figure depicts the bounding distributions and the truncated standard normal distribution at point $(0.355,\ -0.820,\ -0.810,\ 0.218,\ 2.72 \times 10^5)$ which belongs to the estimated set in Table 6.

Table 7: Second Stage Estimation Results: Global Bidders

	Top 50 MSAs	Other MSAs	RSAs
Mean (μ)			
Pop-MHz	0.399	0.365	0.328
Land Area (mi^2)		-16.583	
Small bidder dummy		-5421.774	
Variance (σ^2)			
Pop-MHz	0.023		0.335
ϑ		2.1×10^6	

A confidence interval based on a subsampling method is in progress.

location, the greater the negative effect will be on the mean value of the license. If a license is located one meter further from a bidder's business, the bidder will value the license 66 cents to one dollar less. The negative sign of the coefficient was expected since a local bidder tends to bid more aggressively for a license that covers its own location. It is not clear whether the area has a negative effect or a positive effect on the mean of a license's value from the estimated set. The estimated set also includes zero.

Figure 2 provides the graphical representation of the second stage estimation for local bidders. The figure shows that the truncated normal distribution falls within the bounding distributions at one point in the estimated set.

Table 7 presents the second stage estimation result for global bidders. For global bidders, I used the midpoint in the estimated set of the complementarity coefficients to construct lower and upper bounds of the stand-alone values of a license, instead of the extreme points as in (9) and (10). This is to reduce the computational burden of finding all solutions that minimize the second stage sample criterion function in (14). The coefficients of *geocomple* and *travelcomple* used are 0.0138 and 2.458×10^{-3}, respectively. To capture heterogeneity across the global bidders, I included a small bidder dummy which takes the value of 1 if a global bidder's winning collection contains less than 0.5% of the total Pop-MHz sold. The coefficient estimate for this dummy variable implies that the small global bidders value each license $5,422$ dollars less compared to the other global bidders.

Table 7 shows that global bidders value Pop-MHz more for MSA licenses than RSA licenses. A unit increase in the Pop-MHz of a license increases a global bidder's willingness-to-pay for the license by 40 cents if the license is associated with one of the top 50 MSA markets, whereas only by 32 cents with an RSA market. Assuming that these estimates correspond to the midpoints of the interval estimates for the coefficients, the results indicate that there is horizontal heterogeneity between global and local bidder groups. In general, local bidders have lower valuations than global bidders, except for rural licenses in their region.

Area_l has a negative effect on the mean of the stand-alone value for global bidders. If a CMA area increases by one square mile, global bidders will be willing to pay 16 dollars less on average. The Pop-MHz of a license increases the variance of the stand-alone value of a license; this effect is larger for MSA licenses.

Using the estimation results, I calculate the markups of winning bidders. The markup of a winning bidder measures the difference between the value of the winning collection of licenses and the sum of its winning bids. As I do not recover the exact value of a collection of licenses for a bidder, the exact markups cannot be estimated. Therefore, I consider the expected markups by taking expectation with respect to the stand-alone values of licenses in a winning collection. I require that the surplus from the winning collection is non-negative. Formally, the expected markup ratio Em_i is defined as follows:

$$Em_i \equiv 1 - \frac{\sum_{l \in W_i} p_{i,l}}{EV_i}\ ,\ EV_i = E\left[\widehat{\beta} x_{W_i} + \sum\nolimits_{l \in W_i} v_{i,l} \middle| \widehat{\beta} x_{W_i} + \sum\nolimits_{l \in W_i} v_{i,l} \geq \sum\nolimits_{l \in W_i} p_{i,l}\right] \quad (16)$$

where W_i denotes bidder i's winning collection, x_{W_i} the vector of (geocomple_{W_i}, $\text{travelcomple}_{W_i}$) and $p_{i,l}$ the winning bid for license l.

Note that Em_i in (16) depends on the complementarities coefficients $\widehat{\beta}$ as well as the coefficient estimates that govern the distribution of a license's stand-alone value. Since the distribution of bidder valuations is only set-identified, Em_i is also set-identified. Instead of estimating bidder markups for each coefficient estimate, I use the midpoint of the first stage estimated

Table 8: Expected Mark-ups

	Local Bidders	*Global Bidders*
1st quartile	10.3 %	13.0 %
Median	25.8 %	30.8 %
3rd quartile	56.4 %	62.6 %
St. Dev.	26.0 %	29.3 %

11 out 76 local winners and 4 out of 21 global winners had zero probability that the value of the winning package is greater than their payment. I set their markups as zero.

set for the complementarity coefficients. The coefficients of *geocomple* and *travelcomple* used are 0.0138 and 2.458×10^{-3}, respectively. I also use a point in the second stage estimated set to calculate Em_i for local bidders. The coefficients for Pop-MHz$^e_{il}$, Dist$_{il}$, and Area$_l$ used are 0.355, -0.820 and -0.810, respectively. For the coefficient of Pop-MHz$^e_{il}$ in the standard deviation, 0.218 is used. This particular point is chosen because many points in the estimated set are around this point. Table 8 presents summary statistics on Em_i.

While the estimated markups are relatively high in general, the variation is quite large. While many winning bidders have zero probability that the value of their winning collection is greater than their payment, many bidders have more than 30% markups implying that they only paid 70% of their value for the winning collection. The estimated markups are higher for global bidders. The median for the global bidders is 31% whereas it is 26% for the local bidders. This modest difference in the markups between the two groups indicates that the bidders were horizontally heterogenous because local bidders have higher valuations for rural licenses in their region. It also suggests that the complementarities were not large enough to convert this horizontal heterogeneity to vertical heterogeneity.

There can be several explanations for the relatively high expected markups compared to the bidder markups generally expected in single unit English auctions. First, since Em_i measures the expected markup of bidder i conditional on the value of bidder i's winning collection being greater than the price, this can be imputed to the large variance estimate in the second stage. The large variance in the distribution of stand-alone values of a license implies that each bidder's different business plan was an important determinant of the profitability of each license. The absence of variables that would capture each bidder's different business plan, such as each bidder's market position and existing spectrum holdings in each market, could be a cause of the large variance estimate of the stand-alone value distribution. Another missing factor in the second stage estimation, due to the lack of data, is the competition level of each market. These missing determinants of a license's value could be a source of the large variance of the winners' expected markups.

Second, the high bidder markups are consistent with the fact that the auction prices were below private transaction prices for similar spectrum licenses. For example, in a private transaction, Verizon paid \$2.85 per Pop-MHz for the 10 and 20 MHz licenses in the 1.9 GHz PCS frequency range, which cover in total a population of 73 million people in 22 key markets.[38] It is true that it is hard to compare prices for different licenses in different frequency bands. Different frequency bands imply different service production costs. Furthermore, these private transactions included very important markets such as New York city and Los Angeles. However, the fact that private transaction prices are consistently higher than auction prices suggests that bidders might have shaded their bids down to enjoy their oligopsony market power in the presence of private information.

Third, the number of licenses far exceeded the number of bidders. Because the local bidders valued the licenses that were close to their businesses and the complementarities are not large enough to offset their strong preferences for only a small set of licenses around their location, there were only a few bidders that were interested in each license. The estimation results show that the global bidders valued the licenses across the US. This could induce global bidders to bid for a large package of licenses to enjoy the realized complementarities. However, there were only a few truly nationwide bidders in the auction. Auction theory implies that these large global bidders would have strong incentives to split licenses at low prices rather than triggering intensive competition among themselves.

7 Concluding Remarks

This paper proposes an empirical model and procedure to estimate bidder valuations in FCC spectrum auctions. Given that there is no well-accepted model of bidding that captures the complexity of the auction format, this paper takes an incomplete model approach by specifying the relationship between the observed bidding behavior and the underlying parameters as a correspondence rather than a function. As a result, only a set to which true parameters belong is identified.

The empirical model of this paper starts from two behavioral assumptions that constitute two revealed preference inequalities. They are consistent with rational behavior that should arise towards the end of an auction. I develop an estimation procedure that generates a map from the observed bidding behavior to a set of distributions of bidder valuations consistent

[38] Also, Cingular paid \$1.60 per Pop-MHz to NextWave for 10 MHz licenses in the 1.9 GHz PCS frequency range, which cover in total a population of 83 million people in 34 key markets. In Auction No. 35 held in 2001, Verizon bid \$5.79 per Pop-MHz. The average price per Pop-MHz in the auction was \$4.37.

with the behavioral assumptions. A part of the procedure employs an estimator developed by Pakes, Porter, Ho, and Ishii (2006). I apply the empirical framework to a spectrum auction held in 2006, labeled AWS-1.

Using the estimated distribution of bidder valuations, I first test whether complementarities exist among AWS-1 licenses. As expected, the estimation results reject the hypothesis of no complementarities among AWS-1 licenses. This is consistent with many empirical studies that found synergies among spectrum licenses.

I estimate bidder markups using the estimated distribution of values for a collection of licenses. The results indicate that the expected markups of the winning bidders, conditional on each winner's profit being non-negative, are relatively high compared to single object auctions. I argue that this result is consistent with the following observations: (i) the auction prices were relatively low compared to private transaction prices for similar spectrum licenses and (ii) the number of bidders compared to the number of licenses was small. This result suggests that there were distortionary effects of private information in the auction, and casts doubt about the current auction format's ability to control strategic behavior of big bidders such as nationwide cellular phone carriers.

An incomplete econometric model can achieve a higher level of generality, which is important especially when committing to a particular model as a data generating process is likely to result in specification error. For spectrum auctions, auction theory has not been able to characterize an equilibrium in a general setting. Hence, there is little known about bidders' bidding behavior. Given this challenge, an incomplete model approach may be the only approach available for a researcher vigilant about specification error. However, interpretation of results from a partially identified model is less decisive compared to a point-identified model. Also one may face a practical issue of finding multiple solutions to a minimization problem, especially when the criterion function is not linear in parameters.

The behavioral assumptions of this study may not use all the information available. For example, the assumptions do not use information about a bidder's value on a license if the bidder bids on the license in an early round and remains the standing high bidder until the auction ends. This is because the assumptions do not account for strategic bidding behavior that can alter the equilibrium played and hence ignore noisy information in early rounds. Although it is likely to be true that a bidder's bidding behavior in early rounds contains noisy information about the bidder's value for licenses, the use of bids placed only on a small set of licenses in late rounds reduces the variation that an empirical researcher needs. As a consequence, the interval estimates for the complementarity coefficients are wide, weakening policy conclusions derived from them. Also, lack of data on important determinants of the value of a license for a bidder, such as the competition level in each geographic wireless service market, led to large variance estimates in the distribution of bidder valuations. For future work, I plan to explore the possibilities of tightening the interval estimates on the complementarities by incorporating more information missing from the behavioral assumptions. I also plan to collect data on variables that captures the competition level of each market.

This paper allows for private information on the bidder-license specific values. Private information in this form could be incorporated because (i) it enters the bidder's surplus in a linear fashion and (ii) the revealed preference inequalities and Pakes, Porter, Ho, and Ishii (2006) give rise to implementation of the fixed effect approach.

The econometric model in this paper does not allow for private information in the complementarities. Allowing for this form of private information is extremely challenging because the dimension of private information equals the number of possible combinations, $2^N - 1$, where N denote the number of licenses for sale. There is neither an auction theory nor an econometric technique that enables a researcher to explore this path. One can avoid this curse of dimensionality by allowing private information to enter the bidder's value as a bidder fixed effect multiplied by the observed complementarities. In this case, one needs an estimation strategy that can account for this unobserved structural error entering the value in a non-linear fashion. I leave this challenge as future work.

References

Ausubel, L., Cramton, P., McAfee, P., and McMillan, J., "Synergies in Wireless Telephony: Evidence from the Broadband PCS Auctions," *Journal of Economics and Management Strategy*, 6(3), 497-527, 1997

Ausubel, L. M. and Cramton, P. R.,"Demand Reduction and Inefficiency in Multi-Unit Auctions," working paper, 1998

Ausubel, L. M. and Milgrom, P., "Ascending Auctions with Package Bidding," Working Paper, 2001

Avery, C., "Strategic Jump Bidding in English Auctions", *The Review of Economic Studies*, 65 (2), 185-210, 1998

Bajari, P. and Fox, J., "Measuring the efficiency of an FCC spectrum auction," working paper, 2007

Bajari, P. and Yeo, J., “Auction Design and Tacit Collusion in FCC Spectrum Auctions”, NBER working paper No.14441, 2008

Brusco, S. and Lopomo,G., “Collusion via Signalling in Simultaneous Ascending Bid Auctions with Heterogeneous Objects, with and without Complementarities,” *The Review of Economic Studies*, 69 (2), 407-463, 2002

Bykowsky, M., Cull, R. and Ledyard, J., “Mutually Destructive Bidding: The FCC Auction Design Problem ”, *Journal of Regulatory Economics*, 17 (3), 205-228, 2000

Chapman, J., McAdams, D. and Paarsch, H. (2006): “Bounding Best Response Violations in Discriminatory Auctions with Private Values”, working paper, 2006

Cramton, P. and Schwartz, J., “Collusive Bidding in the FCC spectrum Auctions,” *Contributions to Economic Analysis&Policy*, 1 (1), Article 11, 2002

Donald, S. and Paarsch, H., “Piecewise Pseudo-maximum Likelihood Estimation in Empirical Models of Auctions,” *International Economic Review*, 34 (1), 121-148, 1993

Englmaier, F., Guillen, P., Llorente, L., Onderstal, S. and Sausgruber, R., “The Chopstick Auction: A Study of the Exposure Problem in Multi-Unit Auctions, ” CESifo Working Paper Series No. 1782, 2006

Fox, J., “Semiparametric estimation of multinomial discrete choice models using a subset of choices”, *The Rand Journal of Economics*, 38 (4), 1002–1019, 2007

Guerre, E., Perrigne, I., and Vuong, Q., “Optimal Nonparametric Estimation of First-Price Auctions ”, *Econometrica*, 68 (3), 525-574, 2000

Haile, P. and Tamer, E., “Inference with an Incomplete Model of English Auctions,” *Journal of Political Economy*, 111 (1), 1-51, 2003

Hortacsu, A.,“ Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market,” working paper, 2002

Kastle, J., “Discrete Bids and Empirical Inference in Divisible Good Auctions”, working paper, 2008

Klemperer, P., “What really matters in auction design,” *Journal of Economic Perspective*,16 (1), 169-189, 2000

McAdams, D., “Partial Identification and Testable Restrictions in Multi-Unit Auctions”, *Journal of Econometrics*, September 2008, 74-85

Manski, C. F. and Tamer, E., “Inference on Regressions with Interval Data on a Regressor or Outcome,” *Econometrica*, 70 (2), 519-546, 2002

Milgrom, P. R, “Putting Auction Theory to Work: The Simultaneous Ascending Auction,” *Journal of Political Economy*, 108 (2), 245-272, 2000

Milgrom, P.R., and Weber, R. J., “A Theory of Auctions and Competitive Bidding,” *Econometrica* 50 (5), 1089-1122, 1982

Moreton, P. and Spiller, P., “What’s in the air: interlicense synergies in the Federal Communications Commission’s Broadband Personal Communication Service spectrum auctions,” *Journal of Law and Economics*,41(2),677-725, 1998

Pakes, A., Porter, J., Ho, K., and Ishii, J., “Moment inequalities and their application,” working paper, 2006

Roth, A., “The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics,” *Econometrica* 70 (4), 1341-1378, 2002

Rubinfeld, D., McCabe, M. and Nevo, A., “The Pricing of Academic Journals, ”working paper, 2006

Salant, D.,“Up in the Air: GTE’s Experience in the MTA Auction for Personal Communication Services Licenses,” *Journal of Economics & Management Strategy*, 6 (3), 549 - 572, 1997

Sherstyuk, K., “Simultaneous Ascending Auctions with Common Complementarities,” working paper, 2002

Sherstyuk, K., “On competitive equilibria with common complementarities,” *Mathematical Social Sciences*, 46(1), 55-62, 2003

Yeo, J., “Stylized Facts of FCC spectrum auctions,” working paper, 2008

Zheng, C., “Jump Bidding and Overconcentration in Decentralized Simultaneous Ascending Auctions,” working paper, 2003

Appendix A

Lemma 1 *Assume $w > 0$. Let $f(\cdot)$ and $g(\cdot)$ denote density functions and $*$ the convolution operator. $\int_{-\infty}^{w}(w-z)(f*g)(z)dz \leq \int_{-\infty}^{w_1}(w_1-x)f(x)dx+\int_{-\infty}^{w_2}(w_2-y)g(y)dy$ where $w_1+w_2=w$ with $w_i \geq 0,\ i=1,2$.*

Proof $\int_{-\infty}^{w}(w-z)(f*g)(z)dz = \displaystyle\int_{-\infty}^{\infty}(\max\{w-z\},0)(f*g)(z)dz \leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\max\{w-z\},0)f(x)dx\ g(y)dy \leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\max\{w_1-x,0\}+\max\{w_2-y,0\})f(x)dx\ g(y)dy \leq \int_{-\infty}^{\infty}\max\{w_1-x,0\}f(x)dx+\int_{-\infty}^{\infty}\max\{w_2-y,0\}g(y)dy = \int_{-\infty}^{w_1}\max\{w_1-x,0\}f(x)dx+\int_{-\infty}^{w_2}\max\{w_2-y,0\}g(y)dy$. The first inequality follows from Young's inequality for convolution. The second inequality follows from convexity of the function $max\{x,0\}$: $max\{\alpha x_1+(1-\alpha)x_2,0\} \leq \alpha\max\{x_1,0\}+(1-\alpha)\max\{x_2,0\}$, $0 \leq \alpha \leq 1$ where αx_1 and $(1-\alpha)x_2$ corresponds to w_1-x and w_2-y, respectively. ■

Proof of Proposition 1: I show the first part: given the opponent opening with $b_2(L_2,\mathcal{T}\backslash\{L_2\}) = (0,-\infty,-\infty)$, it's optimal for bidder 1 to open with $b_1(L_1,\mathcal{T}\backslash\{L_1\}) = (0,-\infty,-\infty)$ where L_i is the license that has the highest stand-alone value to bidder i. Suppose $v_{1a} > v_{2b}$. Assume that bidder 1 opens with $b_1(a,\mathcal{T}\backslash\{a\}) = (0,-\infty,-\infty)$. Let $\varphi\ (v_{1a},v_{1b},v_{1c}) \equiv \int_0^{v_1}(v_1-v_2)dF\left(v_2|\, v_{2a} \geq v_{2b}, v_{2a} \geq v_{2c}\right)$ denotes the expected utility from following the strategy in Proposition 0 that will arise if bidder 2 also opens with $b_2(a,\mathcal{T}\backslash\{a\}) = (0,-\infty,-\infty)$. Bidder 1's ex-ante expected surplus from the equilibrium strategy is $q(v_{1a};v_{1b},v_{1c}) \equiv \dfrac{1}{3}\varphi(v_{1a},v_{1b},v_{1c}) + \dfrac{1}{3}\left[v_{1a}+\int_0^{v_{1b}}(v_{1b}-x)dG(x)\right]+\dfrac{1}{3}\left[v_{1a}+\int_0^{v_{1c}}(v_{1c}-x)dG(x)\right]$ where $G(x)$ denotes the conditional distribution of v_{2c} : $G(\cdot) = F(\cdot|v_{2c}\ \leq v_{2b}\ , v_{2a}\ \leq v_{2b}\ )$. By the symmetry, $q(v_{1a};v_{1b},v_{1c}) - q(v_{1b};v_{1a},v_{1c}) \equiv \dfrac{2}{3}v_{1a}+\dfrac{1}{3}\int_0^{v_{1b}}(v_{1b}-x)dG(x)-\left(\dfrac{2}{3}v_{1b}+\dfrac{1}{3}\int_0^{v_{1a}}(v_{1a}-x)dG(x)\right) = r(v_{1a})-r(v_{1b})$ where $r(t) \equiv \dfrac{2}{3}t-\dfrac{1}{3}\int_0^{t}(t-x)dG(x)$. Since $r'(v_{1a}) = \dfrac{2}{3}-\dfrac{1}{3}\left(G(v_{1a})-G(0)\right) > 0$, we conclude $q(v_{1a};v_{1b},v_{1c})-q(v_{1b};v_{1a},v_{1c}) > 0$.

To prove the described strategy is an equilibrium strategy, I show the interim expected utility from following the collusive equilibrium is greater than that of the competitive equilibrium in Proposition 0. Let's assume $L_1 = a$ and $L_2 = b$ without loss of generality. The interim expected utility from the collusive equilibrium outcome is $v_{1a}+\int_0^{v_{1c}}(v_{1c}-x)dG(x)$ while the expected utility from triggering the competitive strategies is $\varphi(v_{1a},v_{1b},v_{1c}) = \int_0^{v_1}(v_1-v_2)dF\left(v_2|\, v_{2a} \geq v_{2b}, v_{2a} \geq v_{2c}\right)$. The reason the expected utility does not depend on the complementarities is that they will be competed away since they are common. Bidder 1 will accept collusion if $\alpha(v_{1a},v_{1c}) \geq \varphi(v_{1a},v_{1b},v_{1c})$. Let $h(\cdot) \equiv f\left(v_{2a}+v_{2b}|\, v_{2b} \geq v_{2a}, v_{2b} \geq v_{2c}\right)$.

Since $\varphi(v_{1a}, v_{1a}, v_{1c}) = \int (\max\{v_1 - z\}, 0)(h * g)(z)dz \leq \int_0^{v_{1a}+v_{1b}} (v_{1a} + v_{1b} - x)h(x)dx + \int_0^{v_{1c}} (v_{1c} - y)g(y)dy$ by Lemma 1, it's enough to check $v_{1a} \geq \int_0^{v_{1a}+v_{1b}} (v_{1a} + v_{1b} - x)h(x)dx$. The right hand side is increasing in v_{1b} and hence it will hold for all $v_{1b} \leq v_{1a}$ if it holds when $v_{1b} = v_{1a}$. Since the right hand side is convex in v_{1a} and $\int_0^{2v_{1a}} (2v_{1a} - x)h(x)dx \leq \int_0^{v_{1a}} (v_{1a} - x)g_1^{(3)}(x)dx + \int_0^{v_{1a}} (v_{1a} - x)g(x)dx$, it's enough to check if

$$1 \geq 2 - \left[\int_0^1 xg_1^{(3)}(x)dx + \int_0^1 xg(x)dx\right] \tag{A1}$$

where $g_1^{(3)}(\cdot)$ denotes the density of the highest order statistics of three independent random draws. Note that $\left[\int_0^1 xg_1^{(3)}(x)dx + \int_0^1 xg(x)dx\right] = \frac{3E(x) + 1 - \gamma}{2}$ where $\gamma = \int_0^1 F(x)^3 dx$. Hence, given the condition $E(x) \geq \frac{1+\gamma}{3}$, (A1) holds. Under the condition, since the interim expected surplus when they have the opportunity to collude is not smaller than the interim expected surplus from triggering the competitive equilibrium, it is optimal to follow the collusive strategies from the first round. ∎

Condition A *For all $\alpha \in [0, 1]$, the following holds:*
$\left[\int_0^1 xg_1(x)dx + \int_0^{1-\alpha} x\widetilde{g}_1(x)dx\right] \geq 1$ *and* $\left[\int_0^1 xg_2(x)dx + \int_0^{1-\alpha} x\widetilde{g}_2(x)dx\right] \geq 1$
where $g_1() = f(v_{2a}|\, v_{2a} \geq v_{2b} + \alpha, v_{2a} \geq v_{2c})$, $\widetilde{g}_1() = f(v_{2b}|\, v_{2a} \geq v_{2b} + \alpha, v_{2a} \geq v_{2c})$, $g_2() = f(v_{1a}|\, v_{1a} = v_{1b} + \alpha, v_{1a} \geq v_{1c})$ *and* $\widetilde{g}_2() = f(v_{1b}|\, v_{1a} = v_{1b} + \alpha, v_{1a} \geq v_{1c})$

Proof of Proposition 2: The proof of Proposition 1 has already shown that given the opponent opening with $b_j(L_j, \mathcal{T}\backslash\{L_j\}) = (0, -\infty, -\infty)$, it is of bidder i's best interest to open with $b_i(L_i, \mathcal{T}\backslash\{L_i\}) = (0, -\infty, -\infty)$. Consider the case in which both bidders open with $b_i(a, \mathcal{T}\backslash\{a\}) = (0, -\infty, -\infty)$, $i = 1, 2$ without loss of generality. In this case, each bidder's type can be denoted as $(v_{ia}, v_{ia} - \alpha_i, v_{ic})$ with $v_{ia} \geq v_{ic}$ and $\alpha_i \geq 0$.

Suppose the bid on a has reached α_1. Suppose that bidder 2 observes that bidder 1 stops bidding on a at the price of α_1. I want to show that under Condition A, each bidder's expected utility from playing the described strategy is greater than or equal to one from triggering the strategy in Proposition 0 given the common information on the opponent's type given $\alpha_2 \geq \alpha_1$. First, bidder 1 must be better off buying b with payment of zero rather than triggering the SEA strategy. By the same argument given in the proof for Proposition 1, it's enough to check

$$1 - \alpha_1 \geq 2 - \alpha_1 - \left[\int_0^1 xg_1(x)dx + \int_0^{1-\alpha_1} x\widetilde{g}_1(x)dx\right] \tag{A2}$$

where $g_1() = f(v_{2a}|\, v_{2a} \geq v_{2b} + \alpha_1, v_{2a} \geq v_{2c})$ and $\widetilde{g}_1() = f(v_{2b}|\, v_{2a} \geq v_{2b} + \alpha_1, v_{2a} \geq v_{2c})$.

Second, bidder 2 must be better off paying α_1 for a rather than triggering the competitive equiliribrium given $v_{1a} = \alpha_1 + v_{1b}$, i.e. $v_{2a} - \alpha_1 \geq \left[\int_0^{2v_{2a}-\alpha_2} (2v_{2a} - \alpha_2 - x) dH(x)\right]$ where $H(2v_{1a} - \alpha_1) = F(2v_{1a} - \alpha_1|\, v_{1a} = v_{1b} + \alpha_1, v_{1a} \geq v_{1c})$. Since the right hand side is convex in v_{2a}, it's enough to check if the inequality holds for type $(1, 1 - \alpha_2, v_{2c})$:

$$1 - \alpha_1 \geq 2 - \alpha_2 - \left[\int_0^1 x g_2(x) dx + \int_0^{1-\alpha_2} x \widetilde{g}_2(x) dx\right] \tag{A3}$$

where $g_2() = f(v_{1a}|\, v_{1a} = v_{1b} + \alpha_1, v_{1a} \geq v_{1c})$ and $\widetilde{g}_2() = f(v_{1b}|\, v_{1a} = v_{1b} + \alpha_1, v_{1a} \geq v_{1c})$. Since the right hand side of (A1) is decreasing in α_2 given $\alpha_2 \geq \alpha_1$, the inequality holds for all $\alpha_2 \geq \alpha_1$ if it holds for $\alpha_2 = \alpha_1$. Under condition A, the above inequalities (**??**) and (A3) are satisfied. ■

Appendix B

Proof of Proposition 3: Note that $\widehat{\overline{\beta}}_q \xrightarrow{P} \overline{\beta}_q$ for each q since $\widehat{\overline{\beta}}_q$ is on the boundary. This is established by PPHI. Since we assume that $\widehat{t}_0 \xrightarrow{P} t_0$ and $\widehat{t}_1 \xrightarrow{P} t_1$, $\widehat{\eta} \xrightarrow{P} \eta$. Also note that $\widehat{h}(t, \gamma, \eta) = \min\left[\widehat{\overline{\phi}}(t; \gamma, \eta), 0\right]^2 + \min\left[\widehat{\underline{\phi}}(t; \gamma, \eta), 0\right]^2$ is continuous in γ for each t. Hence, $Q_{n_s}(\gamma; \eta)$ is continuous in γ. I want to show that $Q_{n_2}(\gamma; \widehat{\eta}_n)$ converges uniformly to $Q(\gamma; \eta)$ in probability on Γ by verifying the sufficient conditions: pointwise convergence and stochastic equicontinuity. I first show that $Q_{n_2}(\gamma; \widehat{\eta}_{n_1}) \to Q(\gamma; \eta)$ in probability pointwise. Note that

$$\begin{aligned} \left|\widehat{Q}_{n_I}(\gamma; \widehat{\eta}_n) - Q(\gamma; \eta)\right| \quad &\leq \quad \frac{1}{n_s}\sum_{k=1}^{n_s} \left|\widehat{h}(t_k, \gamma, \widehat{\eta}_n) - \widehat{h}(t_k, \gamma, \eta)\right| \\ &+ \frac{1}{n_s}\sum_{k=1}^{n_s} \left|\widehat{h}(t_k, \gamma, \eta) - h(t_k, \gamma, \eta)\right| + \left|\frac{1}{n_s}\sum_{k=1}^{n_s} h(t_k, \gamma, \eta) - E_t h(t_k, \gamma, \eta)\right| \end{aligned} \tag{A4}$$

The first term captures the bias from the first stage estimation. Since $\widehat{h}(t_k, \gamma, \eta)$ is continuous in η, and $\widehat{\eta}_{n_1} \xrightarrow{P} \eta$, this term converges to zero in probability as $n_1 \longrightarrow \infty$. The second term is the error from using an empirical distribution. As the empirical distribution converges to the true distribution pointwise (as well as uniformly), the second term converges to zero in probability as $n_2 \longrightarrow \infty$. The last term represents the fact that the inequalities are sampled only asymptotically. Given Assumption 4 (i) and that $Q(\gamma; \eta)$ is finite on Γ, it satisfies a WLLN and hence is $o_p(1)$.

For stochastic equicontinuity,

$$\begin{aligned}\left|\widehat{Q}_{n_I}(\gamma;\widehat{\eta}_n)-\widehat{Q}_{n_I}(\gamma';\widehat{\eta}_n)\right| \;&\le\; \left|\widehat{Q}_{n_I}(\gamma;\widehat{\eta}_n)-\widehat{Q}_{n_I}(\gamma;\eta)\right| \\ &\quad+\left|\widehat{Q}_{n_I}(\gamma;\eta)-\widehat{Q}_{n_I}(\gamma';\eta)\right|+\left|\widehat{Q}_{n_I}(\gamma';\eta)-\widehat{Q}_{n_I}(\gamma';\widehat{\eta}_n)\right|\end{aligned}$$

The first and third term is $o_p(1)$ for any $\gamma,\ \gamma'\in\Gamma$ since $\widehat{Q}_{n_2}(\gamma';\cdot)$ is continuous and $\widehat{\eta}_{n_1}\xrightarrow{P}\eta$. For the second term, the following holds, given continuity of $\widehat{Q}_{n_I}$ in γ and Assumption 3.

$$\lim_{\delta\to 0}\lim_{n_I\to 0}\left[\sup_{\gamma\in\Gamma}\sup_{\gamma'\in\Gamma}\left|\widehat{Q}_{n_I}(\gamma;\eta)-\widehat{Q}_{n_I}(\gamma';\eta)\right|\right]=0$$

Hence, $\left\{\widehat{Q}_j(\gamma;\widehat{\eta}_n)\right\}$ is stochastic equicontinuous. The second part follows from Proposition 5 (b) in Manski and Tamer (2002). ■